1976

Differences In Verbal Abilities In Relation To Age

Lawrence Joseph Hourany

ABSTRACT

The inevitability of biological decline has been used as a model for interpreting intellectual performance in relation to the aging process. It has led to the assumptions that the process of mental development parallels physical growth and decline, that lowered scores by older people necessarily reflect a decline in general performance ability, and that typical age gradients can be taken as at least approximately representative of intellectual decline for any single individual. The purpose of this study was, in part, to test the validity of these assumptions.

The experimental phase of this study attempted to assess the adequacy of certain measures of verbal ability, to examine the influence of flexibility (adaptability to test conditions) as a contributor to performance differences, and to introduce conditions which would enhance the performance of old people. Four measures of ability were used: The Quick Word Test (QWT), the Remote Associates Test (RAT), the Alternate Uses (AU) test, and a new test, the Index of Associative Fluency (IAF). A special characteristic of the latter was that it also produces a non-recipriocal error score.

Two studies were conducted. Subjects in both studies were selected to represent three age groups: Experiment I, young, ages 17-26, N=63; middle, ages 30-59

N=89; old, ages 60-91, N=73. In Experiment II the N values were 71, 86, and 80, respectively, for the three corresponding age groups. All subjects were administered each of the measures. Special instructions were given for the IAF in an effort to produce a learning set and to enhance flexibility. The second experiment was a replication of the first experiment.

The performance gradients for each of the measures used were found to be similar to the results of other research. Using three age groups there was no significant decline for word recognition tasks and a significant decline for response generation tasks. When the samples were divided into seven decade-age groups, the results were the same except for the PWM which now showed a significant decline. There was no interaction between level of ability and age.

The new test developed for the study (the IAF) correlated significantly with an established vocabulary test. It successfully produced a learning set which transferred to other tests for all age groups. It was concluded that the old can learn as well as the young, given appropriate tasks and appropriate learning conditions.

Analysis of the IAF error score indicated that the IAF instructions produced a facilitative change in the strategies of high ability older subjects. The addition of personality measures in Experiment II demonstrated differential relationships between ability and personality variables. It was shown that flexible performance is in part a personality characteristic.

It was concluded that the interpretation of results of gerontological research must take into consideration the particular type of tests being used and the specific age levels comprising the samples; that performance in old age can be modified; and that changes can be made in test conditions in a way that will cause varying results that cannot be accounted for purely in terms of age and ability. It was suggested that many age differences are related more to stylistic factors than to intellective factors. Implications for future research were discussed.

ACKNOWLEDGMENTS

I would like to acknowledge the contributions of my committee members, Drs. D. N. Jackson, T. Seiss, W. Reitz, and V. Vidmar, whose efforts are unmarked but not unappreciated. In particular I would like to thank my supervisor Dr. Norris Schnore, who knew how to be encouraging when my own awkwardness seemed about to defeat me, and for his continued support and guidance. I want to thank my friends Al Segal and Dr. Vince Smeriglio for their thoughtful criticisms. I would like to make special mention of the thorough and penetrating comments made by Dr. Robert Gardner; and for the warmth and consideration he extended throughout numerous consultations. This went far toward reducing the difficulties in "getting it all together".

Appreciation goes, also, to Betty Ledger and to Wendy Leibensperger for their assistance both in typing the paper and for helping me to meet those ever-present deadlines.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

Some controversy has developed over the apparent contradictions found in research concerned with age differences in intellectual performance. The contradictions refer to reports of decline of intelligence in old age as measured by tests of verbal abilities and findings that demonstrate the maintenance of specific verbal abilities well into old age. Although it is accepted by all researchers that loss in adaptive efficiency is apparently a biological inevitability, there is considerable disagreement regarding the age of onset of loss, the shape of the performance gradient, and the relevance and meaningfulness of ability assessment in old age (Botwinick, 1970; Green, 1969).

The resolution of these disagreements will depend on the specific assumptions made concerning native ability, performance (both individual and group), and testing. Three commonly held assumptions are (1) that mental abilities have been adequately identified and measured; (2) that lowered intelligence scores by older people necessarily reflect lessened general performance adequacy; and (3) that typical group age gradients can be taken as approximately representative of intellectual decline for individuals. A deeper analysis of these assumptions shows that they are not tenable.

To begin with, a survey of recent gerontological

literature clearly supports the position that chronological age is not a good indicator of a person's intellectual status, occupational effectiveness, or emotional and social adaptiveness (Botwinick, 1970; Eisdorfer & Lawton, 1973; Kimmel, 1974). To evaluate the current state of knowledge of the aging process, the American Psychological Association (APA) organized the Task Force on Aging. Concerning the alleged loss of intellectual functioning, the Task Force concluded: "Many studies are now showing that the intelligence of older persons as measured is typically underestimated (Eisdorfer & Lawton, 1973, p. ix)." Thus, ambiguities exist and are of growing public and professional concern.

Although there is no single reason underlying the controversies over the assumptions about aging, two distinct general models do seem to account for some of the differences. The first model takes the position that psychological aging is primarily a genetic process that is affected by cumulative biological deterioration or damage to the organism; it is usually termed the biological model or view. The second model defines psychological age changes primarily in terms of experiential factors; it may be termed the developmental model. Of course, the genetic and experiential factors are not independent but remain interactive throughout the life-span (Hunt, 1961, Ch. 2; Tyler, 1965). The difference between the views resides mainly in the relative emphasis each model places on the two factors. As yet, no model of the aging process has unequivocal support (Baltes

& Labouvie, 1973), but discussion of the issues in terms of separate models may clarify some of the problems.

The Biological View

In a primarily biological view, intellectual ability is essentially an innate capacity or potentiality, and as such it is a relatively "fixed" entity (Carroll, 1974; Jensen, 1969; Wechsler, 1958). It is considered to be a "biological reality" (Jensen, 1969), therefore its growth curve is expected to parallel physical development and to follow a gradual, progressive, and inevitable decline beginning in early adulthood (Bromley, 1966; Wechsler, 1958).

In this view, intelligence is considered to be adequately measured by a certain set of tests, and essentially the same tests are used from mid-teens onward (Wechsler, 1958). This set of tests is seen as measuring representative aspects of a global or general, innate, fixed ability, and a single score can be derived and used as an index of performance (Baltes & Labouvie, 1973). Although these tests have not successfully predicted real life achievement they have been used extensively in the prediction of scholastic achievement (Butcher, 1968). Consequently, scholastic performance is considered to be the best, albeit "rather narrow and select," sample of real life useable as a criterion against which intelligent performance can be compared--throughout the life cycle (Jensen, 1969, p. 7). School related performance does decline with age or with distance from the traditional educational model (Crano, 1974, p. 250; Lawton, 1973; McClelland, 1973), thus, in this view,

it is natural to assume that intelligence declines with age.

In the biological view intelligence is a stable, inherited attribute, readily and consistently measureable. Thus, one may restrict the definition of intelligence to that which an intelligence test measures, using as a criterion other, "traditional" intelligence tests (Jensen, 1969, p. 8). This sort of operationalism is the shallowest use of the term in that, in the absence of conceptual understanding and agreement, it permits trivial distinctions to dominate (Carroll, 1974, p. 292). Nor is the circularity of this approach circumvented when it is claimed that "intelligence is a phenotype (Jensen, 1969, p. 17)" since a phenotype is an expression of the genotype as subjected to environmental (physical and social) influences. Moreover, it is agreed that we cannot measure genotypes (Cronbach, 1975; Green, 1974). Thus, the assertions that 80% of the relationship between genotype and phenotype (the heritability ratio) is attributable to the genotype are the weakest sort of speculation. A more conservative approach suggests that the term intelligence be "reserved for the acquired disposition (Cleary, Humphreys, Kendrick, & Wesman, 1975, p. 19)."

The Developmental View

According to the developmental view, chronological age, rather than indexing an essentially inflexible biological process, serves as a general indication of the probable occurrence of certain social, psychological, and biological factors which then interact to produce developmental changes

(Kimmel, 1974; Neugarten, 1973). Experience is particularly emphasized: it is seen to include the occurence of and, sometimes more particularly, the reaction to circumstances such as puberty, menopause, loss of teeth, loss of a mate, chronic or episodic worry, wars, retirement, and so on. Thus, it may be that the events having the greatest influence on psychological development may be those events that compromise the affective life history of the individual (Lowenthal & Chiriboga, 1973). Ceaseless adaptation to these events provides a "continually changing basis within the individual for perceiving and responding to new events (Neugarten, 1973, p. 312)."

In the developmental view, intelligence is considered to be primarily an acquired skill, and as such it is thought to be a relatively stable attribute which develops and differentiates according to the inner motivation, values, and experiential history of the individual (Ferguson, 1954, 1956; Fleishman & Bartlett, 1969). This view considers not only the formation of an ability but its maintenance and effectiveness in relation to other abilities as the product of life experiences (Ferguson, 1965). Accordingly, alterations occur throughout an individual's life. Some experiences lead to relatively temporary changes in a person's behavior, but influences such as education have a more permanent effect (Green, 1969; Hunt, 1961). Therefore, according to the developmental view, intelligence is not a static process but is, instead, a system of progressive adaptations resulting from the

interaction of the organism and the environment (Piaget, 1947, p. 16)." Consequently, the behavioral domain to be measured is considered to extend beyond that measured by typical intelligence tests (Goslin, 1963; Kelley, 1935). The changing nature of the construct is seen to be more fruitfully measured by a different subset of measures at different points in the human life span--measures that reflect the changing needs, values, personality, and environment of the individual (Burt, 1954; Garrett, 1946; Guilford, 1967; Lawton, 1972).

All behavior is influenced by the procedures and structures by which society seeks to perpetuate itself. These procedures take the form of various formal or informal mechanisms which are geared to shaping behavior through selective reinforcements. Thus society promotes those aspects of the range of total possible behaviors that are most useful to society. However, when one feature or part of the social mechanism is missing then what it "produces" is incomplete. That is, when the necessary reinforcers are absent or inadequate the person will develop "inadequately." Individuals not experiencing the full "benefit" of the social mechanisms are not going to exhibit fully developed behaviors. This includes all aspects of behavior including, and perhaps most specifically, those aspects subsumed within the construct of intelligence. It also applies at all points in the life span.

Within the framework of skills that are important to

society are skills that are important to the individual. The effectiveness of a given skill may be related to the interaction of social utility, individual preference and need, experience, and biological efficiency. Thus individuals demonstrate different patterns of skills--rarely is one individual equally good in all skills, nor will all of his skills show the same pattern of development. Based on group data, various skills reach their peak effectiveness at different points in the life span (Bilash & Zubeck, 1960; Blum, Jarvik, & Clark, 1970; Chown, 1961). However, it is an unresolved question whether these peaks are due to different rates of growth or to different life experiences (Botwinick, 1967). This uncertainty also applies to abilities such as sagacity, social adeptness, foresight, and so on, which account for a greater proportion of performance effectiveness in later life (Gilberstadt, 1968; Granick & Friedman, 1967; Trembley & O'Conner, 1966).

In contrast to the more varied patterns exhibited by almost all other skills, verbal abilities show the greatest stability (are the best maintained?) over the life span (Green, 1969; Jones, 1959; Savage & Britton, 1968). This may be because they are the abilities least affected by poor health or because they are the most practiced of all mental abilities (Riegel, 1965). Certainly, mathematical skills, for instance, are less likely to be used regularly by many older persons. While various verbal abilities are not likely to be practiced equally, <u>some</u> practice is universal for community dwelling people. Therefore, to ensure that

the abilities assessed in this research are within the active repertoire of the sampled population only measures of verbal ability will be used.

The stability of verbal skills does not mean they are not subject to fluctuation. As with all skills there are changes (Riegel, 1965), but it is their very stability which make verbal skills so useful for age span research. It may be the one area of common activity for all research participants. Verbal intelligence itself might be seen as a cultural skill reflecting a "bias" of certain cultures (Cole & Bruner, 1971). Thus verbal abilities may be special skills which are required by a culture, and through which intelligence is manifested. The fact that such skills are highly subject to practice and improvement (Fleishman & Bartlett, 1969; Green, 1974) suggests that a developmental view may be a more realistic view of the nature of intelligence than the biological view.

This thesis will contrast the biological and developmental views by providing situations wherein all persons may exhibit performance flexibility and effectiveness (both positive and negative changes might be expected). If performance shifts occur during the experimental conditions, and if older people show improvements commensurate with younger people, then this will be taken as support for the developmental view.

CHAPTER II.

SPECIFIC PROBLEMS IN RESEARCH ON THE AGING PROCESS.

Problems encountered in all research on humans are especially present in studies on aging. Particularly acute are difficulties regarding representativeness of the sample, subject attrition, test sophistication, motivation, and fatigue. Special attention should be paid to these and related problems, some of which are essentially methodological. Other questions which should be asked and which deserve particular consideration in the conduct of research on aging are the following. (These considerations are closely interrelated but they have rather separate consequences.) First, is any test, taken singly or even in conjunction with other measures, appropriate or valid when used with older populations. Some tests attempt to assess functions that are not important in the life of the elderly (Pressey & Pressey, 1967). In that sense, then, the test would be measuring an "irrelevant" skill. Second, does a test measure the same thing throughout the life span? Traditionally, measures have been developed on and for the young and little provision has been made for the effects of adult experiences (Green, 1974; McClelland, 1973). These experiences may be producing essentially incomparable combinations of abilities in older people (Cronbach, 1975;

Green, 1974; Lawton, 1972). The possibility exists that the same test may be measuring different skills at different points in a person's life span. Third, do changes occurring over the life span lead to populations for which age-related comparisons would be misleading? The cumulative effect of life experiences may be such as to lead to unsuspected differences between groups and even among members of various age groups (Hirsh, Kent, & Silverman, 1972). It may be that cross-age comparisons may not be susceptible to detection of these differences despite sophisticated cohort analyses (see, however, Buss, 1973; Schaie, Labouvie, & Baltes, 1973). Fourth, is the measurement situation itself a consistent factor in the assessment of performance of people of different ages? Each of the first three considerations is compounded by the interaction of cognitive and noncognitive factors in assessing performance, and it appears that noncognitive factors assume greater importance in old age (Chown, 1961; Lawton, 1972). It may be that performance effectiveness acquires a larger emotional or personality component over the life span.

I. Test Appropriateness or Validity

In terms of both consistency and relevancy, the validity of a test used with the elderly is a problem primarily because of the lack of established criteria with which to assess ability in old age (Birren, 1964; Carp, 1969a, b; Munnichs, 1965). There are few external referents that can serve as unequivocal criteria against which test performance can be compared.

Perhaps the best measures of ability, both in terms of frequency of use and breadth of research, are omnibus measures of general intelligence. Despite their obvious utility, however, these measures are not without serious challenge (Cleary, et al, 1975; Green, 1974; McClelland, 1973). They were originally developed as estimates of scholastic potential for white, middle-class children. Their great success led somehow to the assumption that they also measured native endowment. Jensen (1969) feels that this is a reasonable occurrence because "the idea of intelligence has justifiably grown considerably beyond its scholastic connotations (p. 8)." He acknowledges that intelligence is not synonymous with "the totality of a person's mental capabilities (p. 19)." But, he is convinced that intelligence is what intelligence tests measure and since they correlate highly with scholastic performance, it is differences in intelligence that produce differences in scholastic performance. Thus, other aspects of mental functioning are subordinate aspects of performance, and even though our society has put a premium on certain (intelligent) kinds of behavior, intelligence "is nevertheless a biological reality (p. 19)." (Jensen might claim that the data from twin studies provide additional, and better, support for a biological view, but much of the data and his interpretations have been strongly repudiated. (Hirsh, 1975; Montagu, 1975) However, the validity of this assumption as applied in a variety of circumstances and with a number of populations (e.g., with second language adults, blacks, the aged, and

others) has been widely questioned (e.g., Cronbach, 1975; Hunt, 1961; McClelland, 1973). The interests, motivations, and experiences of these groups differ from each other in crucial ways (Atchley, 1972; Lawton, 1972). Therefore, when differences in performance occur they are very often traceable to noncognitive factors such as these, and studies which do not take this into consideration may be left with invalid interpretations.

Given the compensatory nature of changing adaptations to life (Lowenthal & Chiriboga, 1973), it is unlikely that even an omnibus measure can provide a satisfactory assessment of life-span cognitive development. However, by using tests that measure those aspects of performance that are over-learned or at least well-learned by almost everyone within a given culture (Cattell, 1971a), the discrepancy between performance and ability would be minimized (Gottesman, Quarterman, & Cohn, 1973; Jarvik & Cohen, 1973), and more realistic and broadly useful appraisals could be obtained (Eisdorfer & Lawton, 1973).

2. Test Consistency or Dimensionality

Consistency refers to the validity of a test in indicating "identical attributes and processes across all segments of the life-span" (Baltes & Labouvie, 1973, p. 175). However, not only may different people use different abilities to accomplish the same task (Guilford, 1967), but the same person may use a different ability or combination of abilities <u>over time</u> (Ferguson, 1965; Fredricksen, 1969), or even at different levels of difficulty for the same type

of task (Guilford, 1968). The interpretive uncertainties generated by this problem are not alleviated by the use of composite scores derived from two or more measures.

The problem of consistency exists even for relatively simple measures. For example, Wechsler (1958) made a distinction between tests that "hold up" well with age and tests that "don't hold". The Digit Span subtest is listed among the latter. Wechsler's data showing a decline in scores on this measure in relation to age have been confirmed by a number of studies (e.g., Taub, 1968; see also Zimmerman & Woo-Sam, 1973). However, Taub has shown that the age deficit is removed when the digits are presented in a manner permitting grouping.

There remains the basic question as to what ability or abilities a test such as Digit Span is measuring and whether this is altered by a change in test format. The test in its altered form may be assessing a somewhat different ability then in its original form. Perhaps this is what occurred in a study by Savage and Bolton (1968). The primary purpose of their study was to investigate the validity of the Wechsler deterioration hypothesis. They used 144 subjects, all of them over the age of 60 and of varying degrees of intactness (from normal to senile dementia). They found, and reported without comment, that the Digit Span sub-test loaded negatively (-.46) on the deterioration factor and had a higher loading than the Vocabulary sub-test (.32 and .26, respectively) on a factor labeled verbal-performance. The Digit Span sub-test is just

one example of a measure that is apparently not a unitary dimension (see Guilford, 1968) when used with all age groups. Consequently, it should not be used as though it represents a single ability or measures the same ability at different ages.

3. Performance Differences and Sample Composition

Many studies may have produced misleading results because of the tendency to combine disparate samples. For example, some studies have included institutionalized subjects with non-institutionalized subjects (e.g., Fisher & Pierce, 1967; Friend & Zuback, 1958; see also Botwinick, 1970) despite research that suggests performance differences for the two populations (Savage & Bolton, 1968; Tyler, 1965). Although some studies have found no significant test score differences between various subgroups at a given age level (Anderson, 1967; Lieberman, 1969), in general, such factors as the specific age grouping used, environment, and occupation introduce systematic biases that alter the representativeness of the sample (Riegel, Riegel, & Meyer, 1968).

Most studies combine their samples according to age groupings that are arbitrary or the result of convenience (see Appendix 1), a practice that tends to reduce the generality of the findings. Studies that do not take into consideration the varied age gradients of different abilities consequently obscure the actual performance relationships. That is, if ability L increases slightly to age 50, ability M is maintained, but ability N declines drastically at

procedure using only a combined score or including 50 year olds with the "young" group or with the "old" group would produce radically different results. For example, a growing portion of recent research on aging has been longitudinal and it has consistently demonstrated performance gradients for individuals which are in sharp contrast to gradients from cross-sectional studies. In addition to finding different age gradients for different abilities (Cattell, 1971a; Schaie, et al., 1973), the evidence has also demonstrated wide variation among performance gradients for specific individuals (Baltes & Labouvie, 1973; Blum, Jarvik, & Clark, 1970). Although some abilities seem to show decline, some remain stable well into old age, and some even increase (Blum, Fosshage, & Jarvik, 1972). For those abilities which show decline, the point at which this decline becomes significant seems to be related to general survival factors or health status (Riegel, et al., 1968). That is, the presence or absence of decline appears related to what is important to the person in terms of life pursuits, and it is related to the biological condition (or the "psychological base" in Cattell's terminology) within which the person must function.

In some studies different ability levels have been combined despite the findings that the level of ability may be more important than age. Chown (1961), for example, found that rigidity was associated more with low intelligence than with age. On the other hand, important research such as Terman's study of gifted people and the Berkeley Growth

Study can trace their results, in part, to the fact that high aptitude subjects were used (Tyler, 1965). Bayley (1968), in a review of her work at Berkeley, felt that many behavioral patterns were obscured as a result of combining groups that did, in fact, have some differences. It has also been found that decrements reported for older groups ceased to exist when above average subjects were used (Keevil-Rogers & Schnore, 1969). Other results also indicate that the more able the individual was initially, the more likely he is to retain and even improve his level of performance (Riegel, 1965; Tyler, 1965). Analysis according to aptitude level should, therefore, be an integral part of age research.

Homogeneity is not completely achieved by restricting groups in regard to age or even aptitude as within any group there may be more or less distinct subgroups in which members behave according to standards which have evolved and become appropriate for them. These "norms set gross limits on appropriateness and relevance...in terms of the survival of the individual and the social structures of which he is a part (Lawton, 1972, p. 124)." While it is true that there are probably as many similarities as differences among people grouped according to age, there may be group qualities which differentially interact with the aging "process" (Hirsch, et al., 1972). This study will attempt to reduce the influence of extraneous group factors by using only white middle-class females.

4. Noncognitive Factors in Performance

It appears that because of the differential operation of noncognitive factors the appraisal of adaptive potential becomes increasingly complex and difficult over the life-span. The interaction of a very large number of noncognitive variables with various aspects of performance has been demonstrated, and these effects appear to be more acute in old age (Brinley, 1965). The interactions of some of these variables or factors and their effects on performance merit specific discussion. These include emotional and motivational factors, the use of familiar tasks, and creativity as a stylistic mode of response.

Emotional and Motivational Factors

Studies have shown that the elderly suffer some performance decrement due to emotional involvement with the test materials, possibly in interaction with attitudinal factors that are independent of the material (Friend & Zubeck, 1958; Reichard, Livson, & Peterson, 1962). Researchers have attempted to reduce emotional reactions and improve motivation and interest by soliciting the cooperation of volunteers (Hulicka, 1967; Keevil-Rogers & Schnore, 1969; Miles, 1965).

As noted earlier, it may be that variables such as motivation and emotional reactions which most researchers try to deal with separately should somehow be included in any appraisals of performance. That is, if a variable (such as reaction to the test conditions) affects performance on a test, then it should be treated as part of that ability (Wechsler, 1958, p. 14; see also review in Butcher, 1968,

pp. 224-6). However, the consensus seems to be, at least in gerontological research (Eisdorfer & Lawton, 1973), that the measurement of ability is made more difficult by the varied influences of emotional and motivational factors and therefore should be treated separately. Nevertheless, while test performance is a more restricted circumstance, it parallels achievement in real life in that it has been shown to result as much from motivational and personality factors as from pure "intelligence" (Butcher, 1968; Furneaux, 1960).

Task Familiarity

Age differences in performance on tests may be due more to an inability to reorganize existing abilities than to a lack of a specific ability. For example, some performance decrement has been shown to be due to unfamiliarity with methods or strategies necessary to high performance on a test, a condition known as method variance. If loss is due simply to unfamiliarity with the strategies necessary to cope with the content of a measure, then the test procedure should include training or retraining in the basic procedures whereby proficiency is obtained: for example, one learns to solve problems after learning how to identify a condition as a "problem" (Birren, 1964).

Also, performance may be poor on unfamiliar tasks because of overlearning on familiar tasks (Ferguson, 1956). Continued experience may lead to a sharpening of a given skill, but it may also limit its variability in application (Guilford, 1967; Hunt, 1961). As a partial solution to this possibility, it has become common practice to interpret performance differences

in relation to tasks which permit a person to use his experiential repertoire rather than requiring him to perform a task in the context of new learning. This practice receives support from studies which have shown no loss or even an improvement for those tasks requiring well-learned responses, such as information, vocabulary, and synonym tests (e.g., Chown, 1961; Jones, 1959; Riegel, 1959; Trembly & O'Connor, 1966), in contrast to studies which required "adaptive learning" (e.g., Arenberg, 1967; Bromley, 1966; Rabbit, 1968; Savage & Bolton, 1968).

In some studies, it has been shown that with the addition of practice conditions, "rest" intervals, or self-pacing, the elderly show improvement with little or no comparable improvement for the young (c.f. Botwinick & Thompson, 1967; Brinley, 1965; Szafran, 1965).

Creativity

Even though creativity is associated with cognitive intelligence, it is also related to noncognitive factors such as intuition and flexibility. Bromley (1965, 1967) attempted to demonstrate that creativity declines with age. He used a measure, the Shaw Test, that requires the arranging of four wooden blocks into a variety of orders, the quantity of responses and "quality" of the order determining the degree of "creativity." Quality was defined as the attainment of a high level concept achieved via the ordering of the four blocks into a specific series according to height, weight, position of a notch, and other characteristics.

He found that the correlation between age and high quality responses was -.52. When intelligence scores were held constant, the correlation dropped to -.19. In another study, (Inglis, 1965) poor performance on the Shaw Test was shown to be related to learning and memory defects.

Is what has been lost creativeness or just one aspect of creativeness, perhaps flexibility? There is no evidence that the Shaw Test is a measure of creativity, although Bromley does suggest that it may be "regarded as an operational measure of creativity" because it is 'open-ended' (1967, p. 33). Bromley also notes that "rational responses...are highly correlated with intelligence, and...reflect creativity (p. 33)." Kogan (1973) in his review of this work says that the first premise "would imply that any test with a divergent structure is a face-valid measure of creativity (p. 149)." He goes on to say that the second statement implies "that one can study one kind of thinking process with tasks that also reflect a very different type of thinking process (p. 149)." Since the age-quality correlation declined to a nonsignificant -.19 when intelligence scores were partialed it is likely that the test measures many functions but is primarily a measure of intelligence.

A crucial component of tasks like the Shaw Test is the set of expectations the subject brings with him to the testing situation (Johnson, 1972, p. 161). When the subject is uncertain as to what to do, he becomes more conservative and tends toward rigid-appearing behavior. This is more pronounced with age. The unfamiliar stimuli and requirements

of the Shaw Test might be expected to produce these results with older people. Very different results were obtained in a study using more familiar, everyday objects. In this study college students and older people (ages 62-85) of approximately equal education were asked to "group the objects into the most comfortable number of categories (Kogan, 1973, p. 168)." The results showed similar modes of cognitive functioning in the two groups. Apparently both groups knew what to do--neither group being "fooled" by the task--and what appears to be operating is general reasoning (Johnson, 1972, p. 221, Horn, 1970). If there is any special ability that is required by the task, most likely it would be cognitive flexibility, which is primarily an attitudinal or stylistic mode of responding (p. 222).

The Biological View: Implications for Specific Research

Adaptive learning has for many years been closely associated with physiological aspects of ability, an association that has had profound influence on research methodology in studies of learning and learning theories. In 1934, Ruch (in Korchin & Basowitz, 1957) hypothesized that age deficits were due to a "lowered biological plasticity" which reduced effectiveness on tasks requiring the reorganization of established patterns of behavior. This notion is embodied in a theory developed by Cattell (1941, 1963; Horn & Cattell, 1967) that relates performance decrements to deterioration in the physiological base upon which intellectual functioning is dependent. Cattell, in

this theory, claims that there are two basic types of intelligence which he has termed fluid (Gf) and crystallized (Gc). Gf refers to "the major measurable outcome of the influence of biological factors on intellectual development," such as heredity and CNS damage, while Gc is "the principal manifestation of a unitariness in experiential-educative-acculturation influence" (Horn & Cattell, 1966). Gc abilities are expected to remain stable since they are considered to be relatively impervious to noncognitive factors. That is, these abilities are relatively unaffected by momentary fluctuations in motivation, interest, or fatigue, by special instructions, or by differences in conceptual style (Green, 1974; Johnson, 1972, p. 296). Both theoretically and on empirical grounds, Gc abilities are seen as pervasive, superordinate factors which determine performance in other areas (Cattell, 1971b, p. 14). However, it has been suggested (Cronbach, 1969) that results derived from measures separated on the basis of "cultural" and "physiological" distinctions may be artifactual and that the researcher may really be dealing with acquired differences for both Gc and Gf.

In reference to a biological orientation, Bromley (1956) claims to have established a connection between Gf and creativity. He goes on to assert that the Shaw Test is a measure of Gf and thus reflects loss in the neuro-physiological base. However, as noted earlier, the Shaw Test seems to be more a measure of Gc than Cf. Thus, the observed declines (which were derived from cross-sectional data)

"may simply represent a by-product of decline in other intellective functions (Kogan, 1973, p. 150)." Rather than reflecting permanent loss, the Shaw Test may be measuring Gc abilities that have fallen into disuse. If so, then proficiency may be revived through re-training: the consequence being an increase in "creativeness".

Actually, recent reviews (see Buss, 1973, p. 473) of Cattell's work indicate that the interpretation of the second-order factors can be reversed: the factor having the highest heretability ratio (largest genetic component) is defined by the salient variables for "crystallized" intelligence. At this time the Gf/Gc theory of intelligence is speculative, and because it overlooks the crucial role of psychological and social elements in behavior, a great deal more refinement is necessary before it's usefulness can be evaluated.

Modifiability of Performance:

The Developmental View

There are good reasons for believing that many contemporary attitudes toward the aged are founded on erroneous premises. For example, in our society "old age has actually become synonymous with age of retirement (Gottesman, et al., 1973, p. 394)." Attitudes such as this may actually be partial causes of the condition known as aging: as with children, when poor performance is expected, perhaps that is what ensues and is then considered a necessary part of the condition (Nardi, 1973). In fact, reviews of recent research "on the health of recent retirees found many more reports of

improved health than of health decrements (Kalish, 1975, p. 108)." Thus, whatever is being labeled "old age" may be more negative myth than reality.

Several factors may be cited in support of the latter contention. First, there is the accumulating evidence that basic competence exists well into old age (Birren, 1970; Eisdorfer & Lawton, 1973). Second, it has been shown that social and psychological factors have a great influence on the aging process (Gottesman, et al., 1973; Neugarten, 1973). Third, many noncognitive factors have been shown to both inhibit and enhance performance (Arenberg, 1973; Brinley, 1965).

Based on such evidence, it is suggested by the <u>developmental view</u> that since behavior does change over time and since everyday events do affect development, then efforts can be made to beneficially influence this process. Previously, the decrement model "led to a reluctance to consider the utility of modification research (Schaie, 1973, p. 152)." This reluctance stemmed from a low opinion, generally held among young and old alike, of the intelligence of older persons, which in turn was a condition perpetuated at least partly by past and present educational disadvantages and by a widespread lack of environmental supports (Eisdorfer & Lawton, 1973). If, as suggested by the developmental model the underlying competence is still present, then what is needed is for older people to recognize the necessary cues in any performance situation (Cole & Bruner, 1971). Under such an interpretation, it is not training or a new skill that is

needed but revitalization and transfer of skills already possessed. The process may have as much to do with personality as with ability factors (Cattell, 1971a; Kalish, 1975, p. 110).

Researchers have not, as yet, succeeded in identifying all the factors influencing performance. This may not even be a feasible objective, especially since none of these factors operate in isolation. For example, efforts to improve performance in older persons by increasing their motivation have not always been successful. Ganzler (1964) conducted a study that failed to demonstrate that specially motivating instructions will elicit a greater improvement in performance for old (ages 65-75) than for young (ages 25-35) men. There were high and low motivation groups, and the experimental manipulations were effective: the high motivation groups did pursue the tasks significantly longer than the low motivation groups. However, although the instructions did change behavior, there was improvement in performance on only some of the tests, and the young and old men improved by approximately the same amount. On some tests, there was no improvement at all. Performance was found to be due, in part, to lack of motivation and in part to not knowing what to do on the test (Birren, 1964). Thus, efforts to improve performance also need to include relevant orientation toward the test situation (Baer, 1972).

Personality and Flexibility: Resistance to Change

Among the more potent and lasting influences on

performance are personality traits (Rokeach, 1971). This may be because the interaction of these traits with cognitive processes operates to form what has been termed cognitive style. These cognitive styles appear to be a central mechanism by which a person interacts with his environment (Kimmel, 1974; Vick & Jackson, 1967). They seem to resemble a "response set" and as such they have been examined in a variety of contexts (e.g., Fredricksen & Messick, 1959; Monge, 1969). There is a variety of styles ranging from a tendency to respond in a socially desirable way to a tendency to take risks in some or many different kinds of circumstances (Jackson, Hourany, & Vidmar, 1972).

It is difficult to ascertain what actually constitutes a cognitive style, but flexibility appears to be a relatively enduring aspect of performance and has been investigated both as an ability and as a personality variable (Chown, 1959, 1961; Guilford, 1967). Regardless of its designation, flexibility exerts a profound influence on performance (Botwinick, 1967).

When flexibility operates as an aspect of personality it has been shown to interfere with performance by producing a "resistance to change" (Chown, 1961; Monge, 1969). It has also been shown to produce performance deficits with age when it functions as a component in cognitive aspects of behavior (Guilford, 1967). In this latter case the testee is expected to <u>spontaneously</u> exhibit flexible behavior. Unfortunately, many of the ability tests used to measure it are unusual, thereby requiring an individual to be what Guilford

has termed "adaptively" flexible before he can be "spontaneously" flexible on the actual test. Therefore, the decrements may not be due so much to a reduction in flexibility as, again, to a lack of relevant test experience.

Various other noncognitive factors influence performance. Harlow (1949) has shown that task efficiency increases slowly until some point at which the performer seems to "learn what it's all about." This would be a cognitive aspect of performance. Any reduction or limitation in this ability to learn may result in a form of rigidity in old age (Botwinick, 1967). Monge (1969) agrees that this apparent reduction in learning ability may be due to the failure of elderly subjects to acquire a set to learn. However, other variables leading to rigidity of performance are noncognitive and might include emotional factors due to error proneness, situational anxiety, and expectations of failure. Harlow suggested that "it remains as a possibility that the resistance of elderly people to new learning may relate more to associated problems of frustration than to learning limitations (1959, p. 512)." Consequently, the rigidity may be a temporary phenomenon capable of being overcome. If typical measurement instruments or assessment situations do not provide for the establishment of a learning set, then this should be adjusted. For example, in a paired-associates learning task, instructions can be provided for the use of mediators plus allowance of time to search for and use them (Hulicka, Sterns, & Grossman, 1967).

The effects of instructional changes have also been

demonstrated for discrimination tasks. In a study involving the perception of line differences, Coppinger and Anthony (1968) provided subjects with instructions on how and when to speed. The results showed a significant increase in response speed without a loss in accuracy. The authors feel that the instructions helped the elderly to overcome their tendency to be cautious. In another study, using subjective time judgments, McNamara (1968) found that by clarifying the task's requirements and the experimenter's intent, differences between the young (ages 18-25, M=22) and the old (ages 65-86, M=78) were abolished.

In summary, it is clear that modifiability of performance can occur through changes in test conditions that have differentially benefited older people. In regard to flexibility, the problems encountered when measuring it as an ability are equaled by the problems encountered when measuring it as a personality variable. At this point it is uncertain whether experimental manipulations produce increased knowledge or increased motivation on the part of the subject: improvement in performance may be a result of either.

It has, however, been shown that performance styles or characteristics (e.g., spontaneous flexibility) may negatively predict some kinds of performance (e.g., clustering in serial anticipation tasks) while correlating positively with other kinds of performance (e.g., on free recall tasks) (Fredricksen, 1969). Thus, performance style may interfere with, rather than facilitate, performance. However, the

longer learning proceeds within a given context, the more likely it is to be a function of stylistic variables and adopted strategies than it is to be a function of conditions inherent in the task (Fredricksen, 1969; Monge, 1969). Therefore, overcoming task constraints via strategy directions (i.e., test instructions) should result in improved performance, or learning, under more extended conditions than is typical in most experimental tasks (Baer, 1972).

Thus, the issue is not whether performance can be modified, but rather in what ways and to what extent. To assess alterations in strategy a measure has been developed which makes provision for flexible responding. In addition, it provides a non-reciprocal error score. This study will therefore include an investigation of the relationship between alterations in strategies and personality variables. The following personality variables were selected: Conformance; Value Orthodoxy; Change; Innovation; Autonomy; Assertiveness; Impulsiveness; and Achievement. (These variables will be further discussed in Chapter IV and Appendix 6.)

CHAPTER III

SPECIFICATION OF HYPOTHESES AND SELECTION OF TEST INSTRUMENTS

This thesis makes three major assumptions. First, it is assumed that learning ability exists relatively undiminished into old age given good health. Therefore, it is predicted that an improved measure of verbal ability can be developed on which the elderly will show no decline relative to younger persons and that this test will produce a learning set, the effects of which will transfer to other measures. Second, it is assumed that the interaction of unique individual traits (such as personality) and ability is a crucial determining factor in performance, especially in relation to flexible performance on tests. Therefore, this study will use a number of ability and personality measures of flexibility to investigate differences in performance in relation to age. Third, it is assumed that some measures are not equally valid for different sub-populations; some tests will have different interrelationships with other tests, depending on the age groups involved.

Hypotheses

A number of specific hypotheses were formulated on the basis of the general hypothetical position outlined in Chapter I and examined in greater detail in Chapter II. The specific hypotheses were as follows:

Hypothesis I-A: that a new test could be developed that would be valid as a measure of verbal ability. Specifically, the new test would show significant correlations with a well established measure of verbal ability across all age groups.

Hypothesis I-B: that there would be no significant performance differences across age groups on this newly developed test. Specifically, the test would show no significant performance decrement with age.

Hypothesis I-C: that the developed test would produce a learning set. Specifically, performance on tests following administration of the test would be superior to that of subjects for whom the test was given after the other tests.

It is believed that in this study noncognitive factors would account for some performance differences, and that these factors would operate differentially across age groups at different levels of ability. Thus, two hypotheses were formulated regarding the role of noncognitive factors:

Hypothesis II-A: that analysis of the non-reciprocal new test error score would show shifts in response patterns in relation to age and ability. Specifically, the test would show that high ability old subjects employ varying strategies, that this modifiability suggests that performance rigidity is not universal for old subjects, and that it is related to ability.

Hypothesis II-B: that noncognitive personality variables will relate differentially to performance (cognitive) variables. In particular, it was expected that

flexible performance is in part a personality characteristic.

It is also predicted that some performance differences related to age are in part determined by the measures that are used. It was expected that the multidimensional nature of some tests leads to unsuspected problems in the interpretation of findings. This suggested an additional hypothesis:

Hypothesis III: that the interrelationships among the tests will vary with age. Specifically, the relationship between a purported measure of creativity and other measures of verbal ability will change across age groups.

Examination of the general and specific characteristics of various tests was undertaken with the hypotheses in mind.

Selection of Tests

A number of tests were chosen as the means of evaluating the hypotheses. The tests were to range over a reasonably wide continuum of associativity (Riegel, Riegel, & Levine, 1966). Each requires either the identification or production of an associate to a stimulus word. While the difficulty level of all of the tests encompasses a wide range, the type of response ranges from relatively straightforward to relatively complex and/or remote associations.

Some broad considerations relating to specific tests were given attention in the process of test selection. A number of tests were examined in terms of two main questions: whether they were valuable as group or individual measures, and whether they coordinated readily with other tests (that is, each test should measure a related but distinct aspect

of verbal ability).

The Quick Word Test

The most widely used measure of general verbal ability has been some form of a vocabulary test (Tyler, 1965). The suitability of such a test stems from its substantial correlations with longer and presumably more valid tests of general intelligence (Borgatta & Corsini, 1960a; Miner, 1961), as well as from its stability when used to test quite old subjects (Kaplan, 1965). The Quick Word Test (QWT, Borgatta & Corsini, 1964) has particular research value because it is a group test that not only correlates well with individually administered tests (e.g., the WAIS, Borgatta & Corsini, 1960b) but also because it maintains its reliability whether timed or untimed (Grotelueschen & Lyons, 1966).

The QWT was selected for use as the "standard" for this study, both because of the simplicity of its format and because of the extensive norms that have been developed for it. The adult form of the test has been shown to yield approximately equal means and distributions with all age groups (Borgatta & Corsini, 1964). Also, it is the test that is most clearly a convergent or "intellective" (Gc) measure (Cattell, 1971a). The QWT was to serve a dual function. First, results from it were to be analyzed as results from the other measures would be--the expectation being that there would be no significant decrement across age. Second, it would reflect the ability dimension--the expectation being that, as the measure having the greatest generality, it most likely would represent "general" ability.

The Alternate Uses Test

There are a number of measures demonstrating stable flexibility scores (Guilford, 1967), but the Alternate Uses (AU) test (revised form of Unusual Uses) has perhaps received the widest use and seems to represent the purest measure (e.g., Chown, 1961; Dixon, 1965; Taft & Rossiter, 1966). Although the test is highly speeded, Guilford (1967) reported high correlations (for young subjects) on timed and untimed versions of the test, with no significant score increments with longer time limits. The performance gradient in relation to age showed an increase to the 40s followed by a moderate, linear decline (Chown, 1961). The AU appears suitable for use with older subjects on the basis of its characteristic of freedom from stress (Dixon, 1965).

In contrast to the QWT, the AU is a measure of divergency, that is, each response to a given stimulus must be different. The AU is used as a measure of an ability index, but it is also expected to be the measure most closely related to flexibility as a personality variable. While the AU is primarily used as a measure of flexibility, it has also been considered to be a measure of originality (Rosenbaum, Arenson, & Panman, 1967). The AU has been used extensively in creativity research (e.g., Guilford, 1967; Hoephner, Guilford, & Bradley, 1970). It was selected for this study both because it is a measure of flexibility and because it requires the production of associates.

The Remote Associates Test

While no absolute differentiation can be made between

intellective and creative abilities, it is those abilities cohering under the general rubric of "creativity" that are seen to suffer most as a result of aging processes (Bromley, 1967; Chown, 1961; Riegel, 1965). Creativity is more than response production generated by the use of simple associations. It requires, at least in part, the production of unusual or remote connections between words, objects, and ideas.

Mednick (1962, p. 221) interpreted the creative process as consisting of "the forming of associative elements into new combinations . ... the more mutually remote the elements of the new combination, the more creative the process or solution." Mednick developed a theory that proposed that creative individuals have uniform probability of producing correct responses regardless of the conceptual remoteness of the stimulus properties. To test this theory, he devised a test called the Remote Associates Test (RAT) wherein an individual is confronted with a number of three-word combinations to which he is to supply a logical associate.

The RAT has been shown to correlate with a variety of factors--for example, awards, ratings by teachers and supervisors, and number of patents--and with other tests (Mednick & Mednick, 1967). However, in such studies no effort was made to control for education and level of intelligence. While no complete analysis has been reported the RAT has been shown to be related to a number of associative measures.

Riegel, Riegel, and Levine (1966) analyzed the RAT in

relation to 14 associativity tasks. Using the Creative Personality scale, which has been validated in previous studies and which correlates significantly with the RAT, they selected high and low creatives. They found support for the general hypothesis that high creatives would produce a larger number of different responses in a free-association task than would low creatives. However, in tasks requiring the production of logical responses, the high creatives exhibited reduced associative variability (are less creative than low creatives). The authors' conclusion was clear: "it is inappropriate to restrict an interpretation of creativity to the most elementary level of associative process (that of numerous and inevitable simple associations). More complex forms of organization and of conceptual structure have to be considered (p. 55)."

If performance on the RAT is related more to logical rather than original responses (as, in fact, is suggested in Mednick's own report, 1962), then RAT performance should correlate highly with measures of convergent, or intellective, rather than divergent (creative) thinking. This was, indeed, the finding of a study by Taft and Rossiter (1966) who factor-analyzed a battery of tests including the RAT. The results showed the RAT loading highest on the same factor as the tests of convergent thinking. Similar results emerged from a study by Jacobson, Elenewski, Lordahl, and Liroff (1968) who found vocabulary test performance to be a better predictor of creativity than RAT performance.

The question, then, is twofold: does the RAT measure

some unique ability such as creativity or intellect while overlapping only minimally with other tests? Or is its contribution to ability assessment more clearly accounted for by other tests and therefore redundant? Despite the claims made both for and against it, the RAT seems particularly sensitive to effects associated with aging. It is a convergent measure that requires flexibility in the subjects. As a convergent measure it should yield results when used with various age groups that are similar to results from vocabulary tests on these same groups. In regard to flexibility it should show an early and progressive decline. On the basis of the literature, it might be expected that the flexibility effects should be greatest and scores should decrease in relation to age.

In contrast with measures which essentially require intellective abilities (e.g., the QWT), the RAT is intended to measure creative ability and was used for that purpose in this study. It is suggested that the RAT is a multidimensional measure, that it is especially sensitive to age-related decrements, and that its correlations with other ability measures and with personality measures will differ for different age groups.

Development of a New Test

Each of the above tests has been assumed, at least by its authors, to measure some specific ability, and each, as noted, does succeed to some degree. However, none makes any provision for alternative ways of achieving a given score. Accomplishing this would require the development of

a test which incorporates a number of special features that take into account the following considerations.

Because of the difficulty of interpreting factorially complex measures, many writers (e.g., Cattell, 1966; Guilford, 1967) advocate the use of relatively simple, unidimensional tests. However, even for simple vocabulary tests, format has been shown to make a difference (Krech, Crutchfield, & Livson, 1969). For example, it has been shown (Guilford, 1967, p. 77), using two different formats, one offering multiple choice alternatives and the other requiring the subject to define the word, that the two versions of the same test loaded on different factors. Different methods of scoring have also resulted in altering the factor content, and may even produce different decline curves (Guilford, 1967, p. 208).

The range or level of performance involved is also an important variable. On the one hand, restrictions in the range of possible responses (the "ceiling" effect) limit the very ablest performers, thereby reducing the discriminative effectiveness of the test. On the other hand, a very broad range of possible responses may produce performance differences not related to the ability purportedly being measured. That is, the differences may come about simply because the difficulty level of the test provides opportunities for some individuals to use abilities not required by a simpler form (Guilford, 1967, p. 209). A number of theorists (e.g., Furneaux, 1960; Hunt, 1961; Jamieson, 1969) contend that assessment of mental processes is incomplete

if the measurement situation makes no provision for use of other than a single response set. These speculations are in contrast to approaches which consider accuracy in measurement and prediction to be achieved only by the use of "pure" measures (see Hunt, 1961, p. 310).

The problems encountered in other areas are even more pronounced when novelty is a factor in ability measurement. The influence of novelty is manifested in three ways. First, there is the novelty of content or materials as used, for example, in concept formation tasks (Wetherick, 1969). Content novelty, in and of itself, may be detrimental to performance of older people (Arenberg, 1968). Second, there is the novelty of method as in tasks requiring the subject to use new techniques in old (familiar) situations. Examples of this are seen in the work on functional fixedness. Third, there is the novelty requisite to producing something entirely new. This can be termed process novelty. It is apparent in creative effort.

The novelty factor can be something of a paradox. It has proven useful in the assessment of high level performance and there is a great deal of evidence indicating that novelty is somehow intimately related to performance decrement in the elderly. Does this mean that old people are no longer capable of high level performance involving novelty? Some writers (e.g., Bromley, 1967; Wechsler, 1958) adhere to this view. But there is suggestive evidence that stimulus novelty and complexity are confounded, and it is this latter variable that accounts for many findings of

performance decline (Brinley, 1965; McNanamy, 1968; Rabbit, 1968). An interactionist orientation would assume that the effects of novelty and complexity are separable and that training, or experience, can act to reduce the influence of either or both. The problem, at this point, is circular and in need of clarification.

The effects of experience, such as predispositions or attitudes, can be altered in the experimental situation by developing in the subjects' minds a set to deal with the material. Set may be induced by training, by instructions, and by manipulations of the task features (e.g., format, content, and so on). Anticipatory or predispositional sets improve performance primarily by speeding the response, but also by alerting the subject to relevent cues or facilitating identification of irrelevant cues (Brinley, 1965). In a multiple choice vocabulary test, for example, the subject has to search for a synonym by analyzing alternative stimuli. Therefore, the instructions should include clues in how to do this.

The effects of set should be particularly pronounced when dealing with flexibility. However, in using flexibility/rigidity as a concept in any behavioral research, especially in aging research, it is important to recognize that flexibility is not a unidimensional factor. It has been shown (Cattell, 1971a) to have at least three major components: (1) classical motor-perceptual rigidity; (2) general flueney (or the lack of it); and (3) ideational flexibility/rigidity (or inertia). The first component involves the perseveration of a motor

or perceptual response; the second reflects the capacity for rapid and/or far ranging retrieval and/or production of responses; the third involves changing or restructuring of the internal or external environment. None of these components has expression restricted entirely to the cognitive domain. In fact, Cattell (1971a) has shown that 90% of the variance in motor-perceptual rigidity is due specifically to personality factors. A general mapping of the rigidity factor has been successfully accomplished (Cattell, 1971a; Chown, 1961). Of paticular note is the research (e.g., Jamieson, 1969; Monge, 1969; Rozeboom, 1967) which has demonstrated that a significant shift in the behavioral expression of this process can be effected by changes in the wording of task instructions.

The Index of Associative Fluency

In order more adequately to cope with the factors of learning set and flexibility, a test, the Index of Associative Fluency (IAF), was developed. This test provides careful instructions which are reiterated periodically during the course of testing, and it permits multiple correct responses. The IAF is a vocabulary test and as such should reflect the continued verbal experiences of the subjects as well as minimize the effects due to content novelty. However, it has some novelty of method, since the test is constructed so as to allow a number of correct alternatives to each stimulus word. Some of the alternative words are in the same grammatical class as the stimulus word, while some are in a different class. Since a number of these

alternatives are equally correct, a high score will depend on the subject's ability to shift set.

The IAF was item-analyzed in a pilot study using a small sample (N=28) and has been shown to have a high positive manifold with no item correlating below .18 (the next lowest correlation is .38; and the range is to .89, with an average r=.64. See Appendix 2 for correlation data. The test's validity as a measure of ability has been demonstrated by a correlation of .81 between IAF raw score and education (in an unusual sample ranging from secretaries to Ph.D. candidates). The highest possible score was 150. However, although every word was selected at least once, the highest score obtained by any subject was 115 (range = 40 to 115).

Reliability coefficients were calculated for the IAF. Eighty-four university students were given the test at the beginning of a class period; the test was readministered six weeks later. The correlation for raw scores was .868 for Form A. Alternate forms reliability was also assessed by giving Form B to twenty-eight students and then giving Form A six weeks later: a coefficient of .721 was obtained. Only Form A was used in this study.

Advantages of the IAF

To obtain a flexibility score and to assess IAF performance more completely, a number of scoring methods have been devised. Since the IAF words represent a wide range of difficulty, from very common to very rare words, the total number correct should be a good measure of what can be termed verbal fluency. Responses coming from different

grammatical classes provide a measure of flexibility.

The IAF also provides an open format in regard to errors. Since the elderly tend to sacrifice speed in favor of accuracy (Botwinick, 1967, p. 172), it is important to use some untimed tests. However, in addition to reduced time strictures it is important to "leave room" for errors. The IAF does this by alerting the testee to the variety of acceptable answers and by providing reasonable alternatives which might entice a guessing strategy. The abandonment of of a cautious orientation, or at least the employment of a less cautious orientation, should be reflected in a non-reciprocal error score.

There are two factors which might produce confounding influences and thereby interfere with IAF performance. One is the subjects' normal tendency to give responses from the same speech category as the stimulus word (Howes, 1967), a tendency which should increase with age (Ervin-Tripp, 1967). The other is the somewhat novel format. It was expected that the instructions and the content would compensate for these influences.

The IAF, finally, was expected to supply a measure of fluency or convergency (raw score), and a measure of flexibility or divergency (number of shifts in grammatical category). It was expected that the format would evoke a special set, the effects of which would transfer to other tasks. It was also expected that the error score would reflect an attitudinal component by its relationship with personality measures, and by a performance shift in relation to ability level.

CHAPTER IV

THE TESTING INSTRUMENTS AND TEST PROCEDURE

In Chapter III, discussion centered on the general properties of the tests used in this study, especially their value in testing the hypotheses. Chapter IV will provide a description of the testing procedure used in the study and the details on each test.

Characteristics of Testing Instruments

Four tests were used throughout the study. The Quick Word Test, Level II (QWT, Borgatta & Corsini, 1964) was used as the measure of generalized verbal ability. It consists of 100 multiple choice items (one 5-letter stem word followed by four 4-letter alternatives) printed on a single answer sheet. Items are arranged in blocks of five with difficulty ranging from low to high within each block and with approximately equal median difficulty for all blocks. Spaces are provided below each alternative for indicating choice. The following is a sample of items:

1.	taper	leer	wick	work	bind
2.	shoot	bang	push	twig	jump
3.	storm	wild	wash	rend	rave
4.	fatal	dire	evil	omen	wish
5.	foray	food	wood	take	raid

Directions: Fill in the answer space for the word that means the same as the first word. If

As is common in age research, age and education are strongly related (negatively). Orthodoxy and Assertiveness also produced low loadings on this factor. Orthodoxy refers to a tendency to have a conservative outlook and to prefer traditional ways of doing things. The predictable relationship between Orthodoxy and age materialized, while that between Orthodoxy and error did not (Botwinick, 1967).

As anticipated in Hypothesis II-B, the data support the prediction that noncognitive and cognitive variables will show differential relationships. The convergency factor (#2) showed virtually no relationship with any personality variable. This supports the expectation that Gc abilities are little affected by noncognitive variables. The emergence of a flexibility factor (#3) supports the notion of independent performance styles. The emergence of a specific IAF correct and error score factor suggests the operation of a special ability or tendency.

Hypothesis III

It was predicted in Hypothesis III that some tests are not suitable measurement devices when used with some age groups. The results of the study did not substantiate this hypothesis.

Specifically, it was expected that the RAT would change in its relationships to other measures for different age groups. Significance tests for the intercorrelations between each of the ability measures (Edwards, 1960) indicates that there are no significant differences between any two measures across age levels except for the RAT.

constructed for use as a measure of both verbal fluency and flexibility. In early use it has been shown to take no longer than the QWT; each takes between 10 and 25 minutes, depending on the age group. The IAF consists of 29 stem words, each followed by 10 words, any number of which may be synonymous with the stem word in some sense. The instructions are very explicit and inform the subject to search carefully but, without making wild guesses, to take a chance when uncertain. He is told to look for words that mean the same as the stem word, and he is also told that many meanings are given. He is further told that when in doubt, he should reverse the comparison. An example:

hard: size firm fluent obdurate franchise
bruise luck plenish arduous difficult.

Here, the answers firm, obdurate, arduous, and difficult would be circled since they all mean hard in one sense or another. Anywhere from three to seven of ten choices should be circled for each item. Every four minutes the subjects are reminded that there is more than one correct alternative for each of the words, to mark all that they have found, and, when encountering difficulty, to reverse the comparison.

A number of different methods are used to score the IAF. The verbal fluency score is simply the total number correct and will be referred to as the raw score. An item is correct if it appears as a synonym in both Webster's New Twentieth Century Dictionary, 2nd Ed. (1968), and Roget's International Thesaurus, 3rd Ed. (1962).

To score the IAF for verbal flexibility, four indices

are used. First, each word is keyed for grammatical class according to the following categories: (1) noun or noun-adjective; (2) verb or adverb; (3) adjective; (4) verb-adverb-adjective or verb-adjective; (5) noun-verb; (6) noun-verb-adjective or noun-verb-adverb. (Reliabilities not computed.)

Second, the IAF words are ranked according to frequency of occurence in reference to the Kucera-Francis (1967) count, and then divided into quintiles (see Appendix 4). The 20% least frequently occurring words receive a weight of five, the next 20% receive a weight of four, and so on, with the 20% most commonly occuring words receiving a weight of one. A single score is computed for each subject using these weights and is referred to as the weighted score. (Reliability=.843.)

Third, each word is given a unit weight of one according to the Kucera-Francis (1967) count. The words are divided into the following three groups: (A) words with a frequency count of two or less; (B) words with a frequency count of more than 2 but less than 45; and (C) words occuring 45 or more times in the Kucera-Francis count. Computations of separate scores are based on the number of times a subject selects a word in a given group. Category A words comprise 4% of the IAF sample. This is a commonly used cutoff point for use in defining infrequency or remoteness (cf. Mednick & Mednick, 1967; Warren & Davis, 1970). Categories B and C are established by dividing the remaining words into two equal groups with 48% of the IAF sample in each. (Reliabilities=.78; .85; and .80.)

Fourth, each word is keyed according to its frequency of selection for this sample. The weights are determined according to the following scheme: five for words selected less than 10% of the time; four for selection frequency between 10-29%; three for selection frequency between 30-49%; two for selection frequency between 50-69%; and one for selection frequency of 70% or more. Using these weights, a single score is computed for each subject. This score is used to determine whether idiosyncratic selection frequencies occur for this sample. (Reliability=.75.)

Procedure: Experiment I

The subjects consisted of 226 female volunteers selected from the following three age groups: young, ages 17-26; middle, ages 30-59; old, ages 60 and over. Subjects in the young age group were selected from undergraduate psychology students (N=62), middle age subjects from YWCA social clubs (N=91), and old age subjects from senior citizens group (N=73). All subjects were informed that the purpose of the study was to develop some new measures of ability which would, hopefully be used in the future to provide more realistic appraisals of abilities. They were further told that their performance would remain strictly confidential and that a summary of the research findings plus their individual results would be sent to them. The students received course credit for their participation. The YWCA and the senior citizens groups were paid a gratuity for each individual participating.

Subjects were tested in groups of 5 to 25, either at

school or, in the case of the middle and old age groups, at their respective meeting places. An effort was made to maintain a relaxed atmosphere in all testing sessions. Each group was randomly assigned to one of the following test sequences: A=1(QWT) 2(AU), 3(IAF), 4(RAT); B=1324; C=2314; D=3124; E=2143; F=3142; G=4321 (see Appendix 5). The only information solicited from the subjects was their age, and highest level of formal education.

Data were analyzed according to the three main age groups initially established, and also by separating the data roughly according to ten year (decade) age groups: 17-26; 30-39; 40-49; 50-59; 60-69; 70-79; and 80+.

Procedure: Experiment II

Subjects in Experiment II were selected according to the same procedures used in Experiment I. They consisted of 237 female volunteers: 71 young students, 86 middle age YWCA social club members, and 80 senior citizens. Experiment II served as a replication study, with two changes. First, two measures were added. The Utility Test (UT, Guilford, 1968) was added because it supplies independent scores for fluency and flexibility. It is very similar to the Alternate Uses test and possesses nearly identical reliability and validity data (Guilford, 1967). In this test the subject has eight minutes to give as many uses as he can for a brick. No requirement is made that each use be different. Two scores are generated, a fluency score (total number of uses) and a flexibility score (number of times the subject shifts categories).

The other measure added was a personality questionnaire, the Adult Personality Schedule (APS). It is comprised of the following eight scales: conformance, value orthodoxy, change, innovation, autonomy, assertiveness, impulsiveness, and achievement (see Appendix 6).

These scales were developed by D. N. Jackson (1967, 1970a, b). They are part of a ten-year effort to incorporate the full force of current knowledge in personality theory and scale construction. Efforts were made to reduce the influence of unwanted response sets (e.g., social desirability and acquiesence), and convergent and discriminant validity standards were adhered to. Moreover, the scales only include items contributing to substantive validity (Jackson, 1970a). That is, if an item was statistically acceptable but failed to provide empirical evidence of consistency with theoretically related items, it was not used.

There were 16 questions comprising each scale, with 8 keyed True and 8 keyed False. The questions are randomly arranged into a questionnaire which the testee has unlimited time to complete. He answers "True" or "False" according to whether he feels the statement does or does not apply to him. For example, to the statement, "In most situations, I usually agree with the opinions of the group," a response of "True" would indicate conformity.

The second change made in Experiment II was in the order of test administration. Only two orders were used: (a) QWT, IAF, AU, RAT, UT, and PAS; and (b) IAF, QWT, AU, RAT, UT, and APS. Two of the groups participating in Experiment

II had also participated in Experiment I. To check whether there were any <u>individual</u> members again taking part in the research it was requested at the beginning of Experiment II test sessions that anyone who had previously participated in Experiment I indicate this on their test forms. Since all subjects were given a code number consisting of an assigned group number, their age, years of education, and two numbers that they had made up themselves, it was a simple matter to compare earlier and later performance. Only six subjects were repeats and no special differences were observed (see Appendix 7).

CHAPTER V

RESULTS AND DISCUSSION

Experiment II of this study was a replication and extension of Experiment I. The means and standard deviations of each of the measures used in both experiments are shown in Table 1. To determine whether the results of the two experiments were comparable, *t* tests were used to compare the two samples. The obtained *t* values for the QWT, RAT, IAF raw and error scores were all non-significant. The *t* value for the AU was 2.41, significant at the .05 level. The mean for Experiment II was significantly higher than the mean for Experiment I. This result appears to be due to the order of testing and not to characteristics of the samples. However, the means for the two samples were within the range of means reported for various samples (Wilson et al., 1960). The age gradients were similar to those found in the general literature (see Schaie & Strother, 1971). Hence, it is concluded that each of the two experiments represents samples from the same population.

To increase the comparability among measures, the scores for each experiment were transformed to distributions having an equal base (M=50) and standard deviation (SD=10). This was done separately for each experiment and for each measure. Except where noted, all table entries will be in these T scores. For both experiments, divisions into

TABLE 1

MEANS AND STANDARD DEVIATIONS ON EACH OF THE SUBSTANTIVE MEASURES FOR EXPERIMENT I (EI) AND EXPERIMENT II (EII)

Age Group	Young		Middle		Old		All	
EI Tests	M	S^2	M	S^2	M	S^2	M	S^2
IAF	63.2	182.3	65.6	339.5	63.8	417.5	64.8	303.7
QWT	53.9	200.6	56.8	323.5	53.9	359.4	55.9	290.5
RAT	16.3	30.7	16.8	33.4	10.4	28.5	14.5	45.3
AU	24.4	49.3	21.1	68.1	13.1	75.3	20.6	90.9
Error	6.5	21.5	7.6	49.8	14.2	116.4	9.2	71.5
N	62 (Mean age=19.3)		91 (Mean age=41.0)		73 (Mean age=72.5)		226 (Mean age=42.6)	
EII Tests								
IAF	65.6	234.3	67.5	362.8	67.6	447.3	67.0	353.6
QWT	55.0	218.6	59.8	296.9	60.3	268.2	58.5	269.0
RAT	13.8	31.3	17.2	32.6	14.1	29.0	15.1	33.4
AU	27.8	61.8	23.1	64.4	17.3	54.9	22.6	78.1
Error	6.5	17.6	7.8	46.6	15.2	161.5	9.9	91.1
N	71 (Mean age=19.3)		86 (Mean age=41.9)		80 (Mean age=74.0)		237 (Mean age=45.8)	

low, medium, and high ability groups were made on the basis of QWT scores (ranges: low=20-45; medium=46-55; high=56-75).

Before proceeding to the main hypotheses, two methodological considerations require examination. First, in regard to the possibility that the age intervals used to form experimental groups might influence the analysis of results, the data were analyzed in terms of three broad age groups and in seven decade-long age groups. With three groups, in both experiments (EI and EII), all variables show a significant main effect for age except the IAF and the QWT. When seven groups (age ranges: 17-26, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89) are used, the results do not change, except for the QWT, for which there is a significant age effect in both EI and EII. The results are presented in Tables 2 and 3. The results are interpreted as presenting evidence for the somewhat greater sensitivity of smaller age groups in detecting small performance variations in relation to age.

The second consideration is in regard to the possibility that level of ability may be an important determinant of performance differences across age. The variance analyses reported in Tables 2 and 3 indicate that there is a significant main effect for ability for every measure. However, none of the measures show a significant interaction between age and ability. It is concluded that the present results do not support the expectation that low ability subjects will exhibit greater performance decrements across age than high ability subjects.

TABLE 2

AGE AND ABILITY VARIANCE ANALYSES: EXPERIMENT I

Variable	Source[a]	3 x 3		7 x 3	
		F	P	F	P
IAF	A	1.19		.86	
	B	15.83	.001	10.41	.001
	AxB	2.16		.69	
QWT	A	1.82		2.95	.01
	B[b]	-		-	
	AxB	1.95		1.08	
RAT	A	58.26	.001	23.62	.001
	B	8.51	.001	15.28	.001
	AxB	.85		1.10	
AU	A	40.01	.001	17.76	.001
	B	8.54	.001	3.88	.001
	AxB	1.32		1.00	
Error	A	25.69	.001	9.25	.001
	B	13.42	.001	7.48	.001
	AxB	1.25		1.20	

Subjects	Ability	Young	Middle	Old
3 x 3 analysis	High	24	32	26
	Medium	24	20	22
	Low	15	39	20

7 x 3 analysis		20s	30s	40s	50s	60s	70s	80s
	High	24	12	15	5	11	11	4
	Medium	24	5	11	4	10	7	5
	Low	15	18	14	5	7	8	5

[a] A=Age; B=Ability; AxB=Interaction

[b] Spurious computation as ability levels were determined using this variable.

TABLE 3

AGE AND ABILITY VARIANCE ANALYSES: EXPERIMENT II

Variable	Source[a]	3 x 3		7 x 3	
		F	P	F	P
IAF	A	.18		1.25	
	B	26.69	.001	24.46	.001
	AxB	1.20		.68	
QWT	A	.51		6.74	.001
	B[b]	-		-	
	AxB	2.19		2.03	
RAT	A	8.41	.001	3.52	.05
	B	36.92	.001	33.35	.001
	AxB	1.09		2.17	
AU	A	43.67	.001	8.44	.001
	B	7.82	.001	7.04	.001
	AxB	1.07		1.35	
Error	A	30.79	.001	8.96	.001
	B	16.83	.001	15.29	.001
	AxB	2.05		.99	

Subjects	Ability	Young	Middle	Old
3 x 3 analysis	High	33	25	24
	Medium	23	31	27
	Low	24	30	29

		20s	30s	40s	50s	60s	70s	80s
7 x 3 analysis	High	33	15	8	4	7	7	10
	Medium	23	15	5	9	10	6	11
	Low	24	9	12	9	10	15	4

[a]A=Age; B=Ability; AxB=Interaction

[b]Spurious computation as ability levels were determined using this variable.

Main Hypotheses

The first hypothesis (I-A) predicted that the IAF would show significant correlations with the QWT for all age groups. The EI correlations of these two tests for the young, middle, and old age groups were .36, .49, and .55; for EII the correlations were .55, .50, and .38. All of the correlations thus differ significantly from chance at the .01 level of confidence. No other measures show as high or as consistent interrelationships. It is concluded that the IAF is an acceptable measure of verbal ability.

The second hypothesis (I-B) predicted that there would be no decrement across age for IAF performance. As shown in Tables 2 and 3, the IAF is the only test for which there is no significant performance effect related to age in any analysis. It is concluded that verbal ability as measured by the IAF is maintained into old age.

Hypothesis I-C predicted that the facilitating effect produced by the IAF would transfer differentially at each age level. To test this assumption, scores for each of the substantive measures were subjected to a 3 by 2 analysis of variance with three age levels and two conditions: (1) when the relevant measure immediately preceded the IAF in the test sequence, and (2) when the measure immediately followed administration of the IAF. As noted earlier, in EII only two test orders were used, and analyses were limited to a QWT-IAF, IAF-QWT comparison and a QWT-AU, IAF-AU comparison. Sample sizes are entered in the Transfer Analyses section of Appendix 5.

For Experiment I, when the QWT was the dependent variable, the main effect of age was non-significant, but there was a highly reliable order effect ($F=14.3$, $df=1/114$, $p<.001$), while the interaction was not significant. No other order effect was significant in EI. For EII the order effect was again significant for the QWT ($F=3.97$, $df=1/228$, $p<.05$) and also for the AU ($F=5.42$, $df=1/228$, $p<.05$).

This may be accepted as evidence that the instructional set produced by the IAF transferred to other tasks. To ensure that the transfer did not work both ways the IAF was used as the dependent variable and the scores subjected to the same 3 x 2 analysis with the QWT first preceding, then following, the IAF. In neither EI nor EII was the order effect significant.

Since the test sequence for EII was always the same, there is no way to determine whether the differential AU performance was due to the facilitative effects of the IAF or to the depressive effects of the QWT. However, since the sequences were counter-balanced in EI, and because the QWT-IAF results are consistent in both experiments, the AU results in EII are accepted as evidence for the operation of positive transfer.

The AU differences for the two samples noted earlier appear, then, to be due to the consistency in the second experiment of the facilitative effects produced by the IAF.

Although positive transfer did occur, the expected age differential transfer did not materialize. The age and

order transactions were non-significant. While the percent change for each age group was not large (around 10% for each age group), the results are consistent with other findings that demonstrate equal benefit for the old and the young (i.e., equal learning).

In order to determine the possible reasons for the successful facilitative effect of the IAF, a number of internal analyses were conducted. The weighted score for the IAF was evaluated for its use as a flexibility measure. The pattern of correlations of the weighted score and for the unweighted score with the other measures was essentially identical (see Appendix 8). The relationships that held for the raw (unweighted) score also held for the weighted score.

To determine whether differential shifts were occurring among the response alternatives on the IAF, the words were divided into grammatical categories. The results show that in no instance was there greater than a 4% difference between young, middle, and old age groups in their word selection patterns (see Appendix 9).

To assess the effects of word rareness, the unit-weight scores derived for the three Kucera-Francis categories were analyzed across ability levels for each main age group. The results are reported in Table 4. The age effects for each category do not reach acceptable levels of statistical significance. However, across each ability level, and for all three categories, there is a highly reliable effect. While this effect is significant for all frequency categories,

TABLE 4

SELECTION TRENDS IN THE IAF SCORED FOR FREQUENCY USING UNIT WEIGHTS BASED ON THE KUCERA-FRANCIS COUNTS

Age[a]	Ability Level			Summary of Variance Analyses			
	Low	Med.	High	Source	F	P	ω^2
Experiment I							
A[b]							
Young	44.0	49.6	53.3	Age	.39		
Middle	42.6	49.2	57.1	Ability	28.57	.001	.20
Old	44.4	52.3	54.3	Interaction	.94		
B[b]							
Young	47.3	51.6	53.7	Age	1.48		
Middle	44.7	50.6	54.8	Ability	18.20	.001	.15
Old	40.9	53.6	50.3	Interaction	1.76		
C[b]							
Young	48.4	47.1	49.1	Age	.79		
Middle	46.8	48.7	54.5	Ability	4.32	.05	.05
Old	45.0	53.9	51.0	Interaction	2.45	.05	.06
Experiment II							
A[b]							
Young	45.5	51.0	53.1	Age	2.54		
Middle	45.5	48.1	54.7	Ability	30.25	.001	.21
Old	47.3	55.3	53.1	Interaction	1.68		
B[b]							
Young	46.3	51.1	57.1	Age	2.34		
Middle	43.1	50.0	57.6	Ability	27.15	.001	.19
Old	41.3	52.3	50.9	Interaction	1.53		
C[b]							
Young	45.4	51.0	52.3	Age	2.46		
Middle	45.3	49.6	54.1	Ability	11.08	.001	.10
Old	44.6	54.6	53.4	Interaction	.56		

[a] EI df=2/220, 4/220; EII df=2/237, 4/237.

[b] KF frequencies: A=0-2; B=3-44; C=45+.

there is a distinct change in the magnitude of the association. The percent of variance accounted for by ability in EI (Hays, 1963) goes from 5% for very common words to 15% for fairly common words and to 20% for rare words. In EII the percentages are 10, 19, and 21. As word rareness increases, ability level accounts for a greater proportion of the score differences. An examination of the means shows that the source of these shifts may be related to two factors. One is the relative homogeneity of the young group. As this is a highly selected sample (college population), the restriction is almost entirely in the lower ranges. The second is the exceptional performance of the high ability older subjects. The high ability group of the old subjects performed as well or better than their young counterparts on this test.

For the sample of old subjects the means are very similar for all three categories with substantial increases across the ability levels. However, for the young in Category C, ability as defined by QWT performance does not predict selection of common words on the IAF and, instead, performance remains fairly stable. The differential findings for young and old result in a significant interaction in EI between age and ability in the selection of common words. This differential effect was not found in EII. The performance of low ability old people in both experiments was approximately one SD below that of the high ability group. Reduced variation for a college sample might be expected on a test constructed by a college person; the test's discriminative validity appears vouchsafed by the range of performance

in the "broader" sample of old people. The reduced variation of the young group, therefore, seems to be due, as noted above, to the homogeneity of the sample.

To evaluate the hypothesis (II-A) that changes in performance are related to age and ability, two analyses were conducted. First, using the same age and ability groupings as were used to evaluate the other hypothesis, the IAF raw and weighted scores for correct and error responses were correlated. For both the correct and the error choices there are no shifts in selection categories and the correlations between raw and weighted scores are essentially unity (average Δ = .98).

In addition to the agreement within groups, there are systematic changes in the means across age. Groups. The expected decrease for error scores across ability levels holds true for all age groups. These results are reported in Table 5. For the correct score the expected increase was obtained for all groups except the high ability old group. This group did not score as well as the medium ability old group. However, this is true only in reference to the raw and weighted total scores. It would appear that the high ability group did not profit from the instructions and were primarily using a cautiousness strategy. This strategy may have been previously acquired, or alternatively evoked by the somewhat opposing instructions "not to omit any possibilities" but to "avoid wild guesses as they will be counted against the score." While this penalty was

TABLE 5

MEANS FOR THE RAW AND WEIGHTED

IAF CORRECT AND ERROR SCORE

Age Group	Ability Level*			Ability Level*		
	1	2	3	1	2	3
	Experiment I					
	Raw Correct Score			Weighted Correct Score		
Young	59.7	65.2	68.0	166.4	188.6	199.2
Middle	55.9	66.2	75.8	155.0	188.6	221.2
Old	55.5	70.7	67.0	155.1	200.0	192.1
	Raw Error Score			Weighted Error Score		
Young	6.9	5.2	3.9	24.1	19.2	14.6
Middle	11.1	6.1	4.8	35.5	21.7	17.1
Old	19.0	12.5	8.4	59.2	41.0	29.4
	Experiment II					
	Raw Correct Score			Weighted Correct Score		
Young	58.1	70.2	75.2	163.8	202.2	220.7
Middle	55.1	65.8	77.4	153.3	186.1	226.5
Old	54.7	74.6	72.1	153.4	217.9	208.9
	Raw Error Score			Weighted Error Score		
Young	7.0	6.8	5.2	25.3	25.8	18.7
Middle	9.8	8.5	5.8	32.6	30.0	21.3
Old	20.6	17.1	9.6	63.5	54.0	32.7

* Ability level: 1 = low, 2 = medium, 3 = high.

never actually assessed, applying the correction to the scores shows that the high ability groups did in fact "outperform" the medium ability groups (EI: 58.2 to 58.6; EII: 57.5 to 62.5). Perhaps the high ability group adopted a maximizing strategy combining both types of instructions.

Considering that the correct scores increase and the error scores decrease, it is evident that improvement in performance for any of the age groups is not a result of random or careless responding. Furthermore, the inverse relation between the correct and error scores for the medium and high ability old groups indicates that the strategy suggested by the IAF instructions was more stringently adopted by the high ability old group. Since only the old subjects show this reversal it is assumed that these subjects differentially benefitted from the instructions. As anticipated in Hypothesis II-A, the data thus suggest that performance rigidity is not universal for old subjects and that it is related to ability.

To appraise the hypothesis (II-B) that noncognitive variables relate differentially to performance, the next analysis included the personality measures. All of the measures were intercorrelated for the total EII sample (N=237) and then subjected to principal components factoring. All factors with eigenvalues greater than 1.0 were retained. The resulting matrix was then rotated to the varimax criterion.

The results are reported in Table 6.* Two things are immediately noticeable. First, except for two variables, a general (actually bipolar) personality factor emerged. This is to be expected in view of the tendency of personality variables to "pull together" in an aptitude-personality factoring (Cattell, 1971a). Second, despite the above, some personality variables contributed to the definition of aptitude-personality factors.

Factor 1 seems best interpreted as a personality factor despite the absence of two variables: Impulsiveness and Achievement. Impulsiveness did not contribute to the definition of any factor, although on the basis of earlier findings (Botwinick, 1967) it was expected to be negatively related to age and possibly to error scores. Achievement proved to be independent of almost all of the personality variables (except Change) and to show no relation to any of the aptitude variables. The remaining personality measures define Factor 1 in a manner consistent with the interrelationships demonstrated in previous research (e.g., Jackson, 1969).

Factor 2 seems to represent a convergency factor in

---

*Note: Cattell (1966, pp. 233-7) suggests that, when measurement reliabilities are not high, factor interpretation be limited to loadings of .40 or higher. However, the general practice is to use loadings of .30 or higher in analyses of this sort (Hoephner, Guilford, & Bradley, 1970). This will be the procedure employed here.

TABLE 6

ROTATED FACTOR MATRIX FOR EXPERIMENT II[a]

Variable	Factor						
	1	2	3	4	5	6	h^2
Age	-25	20	-39	23	-56	03	62
Education	02	13	23	-03	72	06	60
QWT	-01	94	02	-04	10	04	89
RAT	-02	53	36	-02	-14	02	44
AU	17	14	76	-04	26	11	70
UT Fluency	08	15	72	04	16	-01	58
UT Flexibility	11	04	62	-05	22	11	45
IAF	06	50	16	59	14	01	64
IAF Error	-10	-16	-13	95	18	-03	98
Conformance	-54	-05	02	-03	-02	00	29
Orthodoxy	-49	-04	-26	02	-36	06	44
Change	56	-02	33	-02	11	43	61
Innovation	61	-15	12	-14	13	24	50
Autonomy	74	05	04	01	02	-01	55
Assertiveness	40	-09	25	05	43	23	46
Impulsiveness	27	-06	17	-06	03	-18	15
Achievement	05	-04	10	-02	-06	87	77

Note.--Decimals have been omitted for ease of reading.

[a]N=237. Variance accounted for: 70%

the sense of a selective identification function (Cattell, 1971a, p. 429). Each of the ability measures loading this factor requires either the recognition or production of explicitly defined correct responses. The restrictions placed on the response is the essential distinction between these measures and the others employed in this study. The WT most clearly meets the requirements of a convergency measure (Cattell, 1971a; Guilford, 1967) in that the instructions specify, and the format makes provision for, the selection of only one response. Its loading approaches unity on this factor. The RAT and the IAF are, intentionally, not so specifically convergent either through instructions or format and this is reflected in their lower weights. The RAT does require only one response, but the choice is open-ended. The IAF makes specific provision for the acceptability of more than one response. Thus, both the RAT and the IAF permit a variety of responses while remaining essentially measures of convergency.

Factor 3 appears to be a fluency-flexibility factor. It is strongly represented by the two flexibility measures, the AU and UT Flexibility, and in addition is represented by UT Fluency. The AU loading is to be expected in light of its commonly being the best marker for this factor in other research (e.g., Hoephner, et al., 1970). There is some question as to whether to regard the AU as a measure of spontaneous or adaptive flexibility. However, no restrictions or requirements are placed on UT performance and it clearly can be scored for spontaneous flexibility (Guilford,

1968). The inclusion of the UT Fluency score on this factor may be partly artifactual. It has been shown that even with limited time the generation of a large number of responses will increase the likelihood of different or unusual responses (Wednick, 1962; Wallch & Kogan, 1965). Therefore, in most instances when a single measure is scored for both fluency and flexibility, flexibility is a part score and inherently related to fluency. When a correction is made for fluency, the flexibility score becomes fully independent (Clark & Mirels, 1970) and may load on a different factor. The contribution of the RAT loading to the definition of this factor derives from its requirement that the testee produce a response connecting three remotely related words--a task requiring a certain amount of flexibility in juggling disparate ideas. The concordance between the AU and the UT Flexibility variables in combination with the loading for Change indicates that flexible performance is partly attitudinal. The negative loading for age is consistent with extensive research showing a reduction in flexibility with age (Botwinick, 1970; Chown, 1961).

Factor 4 is a specific factor defined by IAF performance. The saturation of the IAF Error score is almost unity and suggests that a very specific ability or tendency is operating. That this is not simply a trivial factor is attested to by the loading for the IAF correct score. That this variable is related to general word power (e.g., GWT performance) and that it produces improvement on related tasks in a transfer paradigm, indicates that a special ability has been tapped or produced by this test.

Factor 5 might be termed an age-education factor.

As is common in age research, age and education are strongly related (negatively). Orthodoxy and Assertiveness also produced low loadings on this factor. Orthodoxy refers to a tendency to have a conservative outlook and to prefer traditional ways of doing things. The predictable relationship between Orthodoxy and age materialized, while that between Orthodoxy and error did not (Botwinick, 1967).

As anticipated in Hypothesis II-B, the data support the prediction that noncognitive and cognitive variables will show differential relationships. The convergency factor (#2) showed virtually no relationship with any personality variable. This supports the expectation that Gc abilities are little affected by noncognitive variables. The emergence of a flexibility factor (#3) supports the notion of independent performance styles. The emergence of a specific IAF correct and error score factor suggests the operation of a special ability or tendency.

Hypothesis III

It was predicted in Hypothesis III that some tests are not suitable measurement devices when used with some age groups. The results of the study did not substantiate this hypothesis.

Specifically, it was expected that the RAT would change in its relationships to other measures for different age groups. Significance tests for the intercorrelations between each of the ability measures (Edwards, 1960) indicates that there are no significant differences between any two measures across age levels except for the RAT.

However, even for this measure the intercorrelations with the other measures differed significantly only for the middle age groups between experiments--a difference that was not predicted. These results are presented in Table 7. Each of the EI and EII differences are significant at a point beyond the .05 level (.32-.02 difference, t=2.08; .61-.17 difference, t=2.83; .45-.06 difference, t=3.58).

The expected change in intertest relationships was not found. Instead, the relationships were quite stable and, except for the EI middle age group, the correlations were of the same general magnitude. Inspection of the means (see Table 1, p. 53) provides no explanation, as performance level was within the range produced by the subsamples. However, an examination of the test orders (see Appendix 5) showed that only one group received the RAT first in the test sequence. The scores for this group were tabulated and it was found that the mean (M=12.8) was over one standard deviation lower than for the remainder of the group. No other differences of this type were found. No explanation is readily apparent for these findings. Although test order seemed not to affect any other test, it may be that RAT performance may be subject to a "warm-up" effect (Guilford, 1967), especially for subjects who are not test sophisticated. However, this surmise remains to be tested. It is concluded that the correlations for the EI middle age sample are due to sampling fluctuations, and the hypothesis of differential relationships between the RAT and other measures has not been supported.

TABLE 7

CORRELATIONS OF THE REMOTE ASSOCIATES TEST WITH EACH OF THE OTHER MEASURES USED IN EXPERIMENTS I AND II

Ability Measures	Age Groups					
	Young		Middle		Old	
	EI	EII	EI	EII	EI	EII
IAF	.26	.54	.02	.32	.28	.24
QWT	.35	.56	.17	.61	.43	.31
AU	.33	.30	.06	.45	.29	.35
N	63	71	91	86	73	80

CHAPTER VI

SUMMARY AND CONCLUSIONS

Although the hypotheses in this study were for the most part supported by the results, it should be emphasized that this was a single sex study and that the conclusions are in reference to the female population only. However, the measures employed were reliable and the samples for each experiment were representative of the population from which they were drawn.

It was anticipated that age grouping might contribute to findings of age differences. This was examined by comparing results derived from use of three broad age groups to results obtained from use of smaller decade age groups. It was shown (for the QWT) that differences among age groups not detected by 3-group analyses did show significant differences when the relatively smaller decade age groupings were used.

To test the expectation that ability level might be a crucial variable affecting performance differentially at various age levels, the subjects were separated into three ability levels on the basis of their QWT performance. None of the measures in either sample showed a significant interaction between ability and age, and this remained true whether three or seven age groups were used. It is concluded that for the measures used in this study, and keeping in

mind the particular manner used in determining various ability levels, there are no differential age effects caused by varying ability levels.

Findings Related to Hypotheses

The findings of the study may be summarized with reference to each of the hypotheses.

Hypothesis I-A, that the IAF would relate significantly to verbal ability, was supported.

Hypothesis I-B, that there would be no performance decrement across age, was also supported. The confirmation of both of these hypotheses indicates that the IAF represented a valid measure of verbal ability that may be especially appropriate for use in age research.

Hypothesis I-C, that the IAF would produce a facilitating effect differentially transferable across age levels, was supported. It was shown that the instrumental set produced by the IAF resulted in positive transfer, not only to a similar task (the QWT) but also to a task of quite dissimilar construction (the AU). Analysis of IAF test scores revealed that it functioned as a differentiating measurement device while apparently producing a set to respond to a variety of acceptable alternatives. The IAF test emphasis on being aware of many acceptable alternatives appeared to contribute to a learning increment for the older group at least equal to that of the young group. It is concluded that some performance decrement exhibited by older persons can be related to learning set deficiencies and that appropriate instruction can operate to reduce that deficiency.

Hypothesis II-A, that noncognitive variables contribute to performance differences across age, was supported. Analysis of IAF error scores showed that the possible adverse test effects due to a cautious strategy on the part of the elderly could be overcome with apparent benefit. Any increase of test errors for old subjects was also accompanied by a trend toward independence (a trend that turned out to be a function of ability level) between the error score and the correct score. Since the IAF scores showed no decrement with increases in age, it is concluded that the elimination of a cautious strategy produces differential benefit across ability levels for older subjects. It is further concluded that this change reflects the adoption of a learning set. Therefore, rigidity is seen as a temporary phenomenon whose effects can be modified or moderated by the introduction of appropriate learning conditions.

Factor analysis of ability and personality measures made in this study demonstrated the existence of factors bearing close resemblance to those generated in other research. Two ability factors, interpreted as a convergency and a flexibility factor, were isolated. In regard to the personality measures, the results corroborated findings from other research. It is particularly noteworthy that the flexibility factor was found in this study to be determined by ability and personality variables, thus indicating that flexibility is to some extent an attitudinal factor. The implications of this finding for future research and educational practice are significant. Changes

in attitudes may or may not be easier to bring about than changes in abilities, but the effects do seem to be more pervasive and long lasting (Botwinick & Thompson, 1967; Foa, 1968; Rokeach, 1971).

Hypothesis II-B, that personality and performance variables would differentially interrelate, was supported. In particular, the results show that, as a performance characteristic, flexibility is negatively related to age while as a personality characteristic it is relatively independent of age.

Hypothesis III, that the tests used would relate differentially for different age groups, was not fully supported. It was found that the relationship of the RAT to the other measures was significantly different for the middle age group. However, these differences were attributed to random error, and the overall relationships were otherwise interpreted as stable for all measures. When the factor analytic results were considered, it appears that the RAT relates most to measures of convergency and only secondarily to measures of divergency. It is concluded that consideration of the RAT as a measure of creativity should be with caution, especially in samples including older subjects. The RAT performance appeared to be influenced by what might be termed a warm-up effect, which should be taken into special consideration when the RAT is used with testees who are not test sophisticated.

The relationship between Education and Assertiveness (Factor 5) is especially worth noting. A person's

performance in a test situation may be partly determined by personality characteristics such as assertiveness or aggression. These characteristics operate in a motivational fashion by producing a willingness to deal with or "attack" problems. But education has been shown to be a potent demographic variable in a variety of researches that show its influence on attitudes. Accordingly, although no causal relationship can be substantiated by a factor analysis, the possibility exists that education might have as much to do with changing a person's attitudes as it has to do with improving his abilities. On the other hand it may be that assertiveness or aggressiveness propels people to pursue education, and education in turn causes higher performance. Either way, it seems highly plausible that short-term efforts to influence personality (attitudes) may have temporary or even long-term effects on performance. (However, see Appendix 10, and Appendix 11.)

The general support for the hypotheses indicates that the developmental model rather than the biological model provides a more reasonable explanation both for the differences in performance over age reported here and for the changes reported in similar research. This model permits a clearer appraisal of the interactions between <u>time</u> (experience, training, and so on) and <u>age</u>. It points the way for more direct analyses of differences to be made. And it at least indirectly suggests ways the aging process can be influenced. Decline in abilities may not be a reversable phenomenon, and certainly it characterizes an increasing

number of people over the age span, however, each person may not show early or even gradual decline. Thus, at any age and for any individual, the possibility for compensation and improvement in performance exists at later ages than is typically recognized. Certainly the age at which "decline" may be "setting in" must be later in life than is commonly expected (Baltes & Schaie, 1974). In fact, the results support Birren (1970, p. 125) who, in reiterating his 1958 position, said "there is the possibility that the psychological norm for the species is one of little change in intellectual functions in the years after 65, given good health."

Future Research

The question arises as to what extent areas of human experience such as creative abilities or moral or social intelligence can be identified as constructs to be measured and analyzed mathematically. But closer correspondence between real-life and test performance might be obtained if the lives of older people were actually observed. For example, it is reasonable to assume that multiple uses of common objects as an activity declines with age. It is not an activity older people engage in or are prepared to engage in. Therefore the fluent expression of varied possibilities for items that have acquired strong associative bondings is unlikely when this expression is expected to be exhibited spontaneously. (Spontaneous expression of multiple uses underlies the interpretation of the AU test.) However, this does not rule out the possibility that this tendency can be enhanced, perhaps even differentially since

there is more "room" for improvement in those who are not currently involved in this or similar activities. Therefore, it is suggested that future research attempt to determine the individual's involvement in daily experience, to relate this to performance adequacy, and to provide a measurement situation wherein stylistic variables are permitted to moderate performance efforts.

Total assessment would, therefore, include a variety of factors not assessed by the usual measuring instruments. For example, capacity for leadership might require the trait of patience that is perhaps more likely to be possessed by a relatively older person, a trait more related to the development of sagacity, and which does not necessarily depend on sheer speed, energy, or dynamic performance in measurable dimensions. The quality of justice dispensed by an elderly judge will not necessarily depend on his IQ or even on his knowledge of the law. Thus, "decline" in one attribute or ability may be compensated by increase in another. And the amorphous processes by which an individual deals with his environment may fluctuate to meet personal needs. On the other hand, the expectations of the culture might produce many of the apparent declines attributed to old age; if a 35 year-old man were forced to retire and play the role of a 65 year-old man, then he too might quickly begin to exhibit signs of "aging."

While dramatic changes are unlikely, elderly people do retain the ability to change, and, more specifically, to improve their performance. In fact, it may be that

stereotypes regarding aging are producing a self-fulfilling prophecy for both the general public and for some professionals. A concerted effort is needed to allay these stereotypes and to find ways to make life for the elderly more enjoyable and effective.

REFERENCES

Anderson, Nancy N. Effects of institutionalization on self-esteem. Journal of Gerontology, 22: 313-7, 1967.

Arenberg, D. Regression analysis of verbal learning on adult age at two anticipation intervals. Journal of Gerontology, 20: 41-9, 1967.

Arenberg, D. Concept problem solving in young and old adults. Journal of Gerontology, 23: 279-82, 1968.

Arenberg, D. Cognition and aging. In C. Eisdorfer and M. P. Lawton (Eds.), pp. 74-97, 1973.

Atchley, R. C: The Social Forces in Later Life. Belmont: Wadsworth, 1972.

Baer, P. E. Cognitive changes in aging: Competence and incompetence. In C. M. Gaitz (Ed.), Aging and the Brain. New York: Plenum Press, 1972.

Baltes, P. B. & Labouvie, G. V. Adult development of intellectual performance. In C. Eisdorfer and M. P. Lawton (Eds.), pp. 157-219, 1973.

Baltes, P. E. & Schaie, K. W. The myth of the twilight years. Psychology Today, pp. 35-40, March, 1974.

Bayley, Nancy. Behavioral correlates of mental growth: Birth to thirty-six years. American Psychologist, 23: 1-17, 1968.

Bilash, I. & Zubeck, J. P. The effects of age on factorially "pure" mental abilities. Journal of Gerontology, 15: 172-82, 1960.

Birren, J. E. The Psychology of Aging. New Jersey: Prentice-Hall, 1964.

Birren, J. E. Toward an experiential psychology of aging. American Psychologist, 25: 124-135, 1970.

Blum, J. E., Jarvik, L. F. & Clark, T. Rate of change on selective tasks of intelligence: a 20 year longitudinal study of aging. Journal of Gerontology, 25, 171-176, 1970.

Blum, J. E., Fossage, J. L., & Jarvik, L. F. Intellectual changes and sex differences in octogenarians: A twenty-year longitudinal study of aging. Developmental Psychology, 7: 178-187, 1972.

Borgatta, E. F. and Corsini, R. J. The Quick Word Test. Journal of Educational Research, 51: 15-7, 1960 (a).

Borgatta, E. F. and Corsini, R. J. The Quick Word Test (QWT) and the WAIS. Psychological Reports, 6: 201, 1960 (b).

Borgatta, E. F. and Corsini, R. J. Quick Word Test. New York: Harcourt, Brace and World, 1964.

Botwinick, J. Cognitive Processes in Maturity and Old Age. New York: Springer Publisher Co., 1967.

Botwinick, J. Gerophychology. Annual Review of Psychology. Palo Alto, California: Annual Reviews, 1970, pp. 239-272.

Botwinick, J. & Thompson, L. W. Practice of speeded response in relation to age, sex and set. Journal of Gerontology, 22: 72-6, 1967.

Brinley, J. F. Cognitive sets, speed and accuracy of performance in the elderly. In A. T. Welford and J. E. Birren (Eds.), Chapter 8, 1965.

Bromley, D. B. Some experimental tests of the effect of age on creative intellectual output. Journal of Gerontology, 11: 74-82, 1956.

Bromley, D. B. The Psychology of Human Aging. Baltimore: Penguin Books, 1966.

Bromley, D. B. Age and sex differences in the serial production of creative conceptual responses. Journal of Gerontology, 22: 32042, 1967.

Burt, C. The differentiation of intellectual ability. British Journal of Educational Psychology, 24: 76-90, 1954.

Buss, A. R. An extension of developmental models. Psychological Bulletin, 80: 466-79, 1973.

Butcher, H. J. Human Intelligence: Its nature and assessment. London: Methuen, 1968.

Cameron, P. Age parameters of young adult, middle-aged, old, and aged. Journal of Gerontology, 24: 201-3, 1969.

Carp, Frances M. Senility or garden variety maladjustment. Journal of Gerontology, 24: 203-8, 1969 (a).

Carp, Frances M. Compound criteria in gerontological research. Journal of Gerontology, 24: 341-7, 1969 (b).

Carroll, J. B. The aptitude-achievement distinction. In D. R. Green (Ed.), pp. 53-78, 286-303, 343-4, 1974.

Cattell, R. B. Some theoretical issues in adult intelligence testing. Psychological Bulletin, 38: 592[illegible]6, 1941.

Cattell, R. B. Theory of fluid and crystallized intelligence: A critical experiment. Journal of Educational Psychology, 54: 1-22, 1963.

Cattell, R. B. (Ed.) Handbook of Multivariate Experimental Psychology. Chicago: Rand McNally, 1966.

Cattell, R. B. Abilities: Their Structure, Growth and Action. New York: Houghton Mifflin, 1971 (a).

Cattell, R. B. The structure of intelligence in relation to the nature-nuture controversy. In R. Cancro (Ed.), Intelligence. New York: Gruneand Stratton, Chap.1, 1971b.

Chown, Sheila M. Rigidity--a flexible concept. Psychological Bulletin, 56: 195-62, 1959.

Chown, Sheila. Age and the rigidities. Journal of Gerontology, 16: 353-62, 1961.

Clark, P. M. & Mirels, H. L. Fluency as a pervasive element in the measurement of creativity. Journal of Educational Measurement, 7: 83-86, 1970.

Cleary, T. A., Humphreys, L. G., Kendrick, S. A. & Wesman, A. Educational uses of tests with disadvantaged students. American Psychologist, 30: 15-41, 1975.

Cole, M. & Bruner, J. S. Cultural differences and inferences about psychological processes. American Psychologist, 26: 867-76, 1971.

Coppinger, N. W. & Anthony, N. The effect of instructions and response demand on the perception of line differences in adults. Journal of Gerontology, 21: 51-9, 1968.

Crano, W. D. Causal analysis of the effects of SES and initial intellectual endowment on patterns of cognitive development and academic achievement. In D. R. Green (Ed.), Chap.7, 1974.

Cronbach, L. J. Heredity, environment and education policy. Harvard Educational Review, Reprint #2, pp. 190-9, 1969.

Cronbach, L. J. Five decades of public controversy over mental testing. American Psychologist, 30: 1-14, 1975.

Dixon, J. C. Cognitive structure in senile conditions with some suggestions for developing a brief screening test of mental status. Journal of Gerontology, 22: 166-73, 1965.

Edwards, A. L. Experimental Design in Psychological Research. New York: Holt, Rinehart & Winston, 1960.

Eisdorfer, C. & Lawton, M. P. (Eds.) The Psychology of Adult Development and Aging. Washington D.C.: American Psychological Association, 1973.

Ervin-Tripp, Susan M. Changes with age in the verbal determinants of word-association. In L. A. Jakobovitz and M. S. Miron (Eds.) Readings in the psychology of language. New Jersey: Prentice-Hall, 1967.

Ferguson, G. A. On learning and human ability. Canadian Journal of Psychology, 8: 95-112, 1954.

Ferguson, G. A. On transfer and the abilities of man. Canadian Journal of Psychology, 10: 121-131, 1956.

Ferguson, G. A. Human abilities. Annual Review of Psychology, 16: 39-62, 1965.

Fisher, J. & Pierce, R. C. Dimensions of intellectual functioning in the aged. Journal of Gerontology, 22: 166-73, 1967.

Fleishman, E. A., & Bartlett, C. J. Human abilities. Annual Review of Psychology, 20: 349-80, 1969.

Foa, U. G. Three kinds of behavioral changes. Psychological Bulletin, 70: 460-473, 1968.

Fredricksen, C. H. Abilities, transfer, and information retrieval in verbal learning. Multivariate Behavioral Research Monographs, No. 69-2, 1969.

Fredricksen, N., & Messick, S. Response set as a measure of personality. Educational and Psychological Measurement, 19: 137-55, 1959.

Friend, Celia M. & Zubeck, J. P. The effects of age on critical thinking ability. Journal of Gerontology, 13: 407-13, 1958.

Furneaux, W. D. Intellectual abilities and problem solving behavior. In Eysenck (Ed.) Handbook of Abnormal Psychology. New York: Basic Books, 1960.

Ganzler, H. Motivation as a factor in psychological deficit of aging. Journal of Gerontology, 19: 435-9, 1964.

Garrett, H. E. A developmental theory of intelligence. American Psychologist, 1: 372-78, 1946.

Gilberstadt, H. Relationships among scores of tests suitabl for the assessment of adjustment and intellectual functioning. Journal of Gerontology, 23: 483-7, 1968.

Goslin, D. A. The Search for Ability. New York: Russell Sage Foundation, 1963.

Gottesman, L. E., Quarterman, C. E., & Cohn, G. M. Psychosocial treatment of the aged. In C. Eisdorfer an M. P. Lawton (Eds.), pp. 378-427, 1973.

Granick, S., & Friedman, A. S. The effect of education on the decline of psychometric test performance with age. Journal of Gerontology, 22: 191-5, 1967.

Green, R. F.(Ed.) The Aptitude Achievement Distinction. Monterey, California: McGraw-Hill, 1974.

Green, R. F. Age-intelligence relationship between ages sixteen and sixty-four: A rising trend. Developmental Psychology, 1: 618-27, 1969.

Grotelueschen, A., & Lyons, J. J. Quick Word Test validity with adults. Psychological Reports, 20: 488-490, 1966.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Guilford, J. P. The Structure of Intelligence. In D. K. Whitla (Ed.) Handbook of measurement and assessment in behavioral sciences. Don Mills, Ont: Addison-Wesley, 1968.

Harlow, H. F. The formation of learning sets. Psychological Review, 56: 51-65, 1949.

Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), Psychology: A study of a science. New York: McGraw-Hill, 1959.

Hays, W. L. Statistics for psychologists. New York: Holt, Rinehart & Winston, 1963.

Hirsch, C., Kent, D. P. & Silverman, S. L. Homogeneity and heterogeneity among low-income negro and white aged. In D. P. Kent, R. Kastenbaum, and S. Sherwood (Eds.), Chap. 27, 1972.

plug:	form	depart	expose	irritate	nag
	goad	goat	chew	block	boost
range:	mix	vamp	rigid	rank	enact
	print	align	drama	home	solve
report:	acquit	opaque	sound	answer	critique
	award	recite	transfer	notice	rumor
set:	fix	class	etch	obscure	inform
	end	assign	trend	impose	clique
shoot:	diverge	recoil	dart	grow	leap
	jail	execute	project	ricochet	emit
take:	bestow	entail	exude	convey	infect
	eat	catch	deduce	deem	steal
tear:	mend	drop	rip	cleave	rage
	rive	rave	raze	ramp	hewn
unite:	sunder	identify	portion	join	discorp
	converge	votive	meet	own	wed
upset:	discompose	decay	revolution	quell	capsize
	denial	oppose	nerve	distress	sense

Hirsch, J. Jensenism: The bankruptcy of "science" without scholarship. Educational Theory, 25: 3-27, 1975.

Hoepfner, R., Guilford, J. P. & Bradley, P. A. Information-transformation abilities. Educational and Psychological Measurement, 30: 785-802, 1970.

Horn, J. L. Intelligence--Why it grows, why it declines. In E. L. Kintz and J. L. Bruning (Eds.), Research in Psychology, Glenview, Illinois: Scott, Foresman, 1970.

Horn, J. L. & Bramble, W. J. Second-order ability structure revealed in rights and wrongs scores. Journal of Educational Psychology, 58: 115-22, 1967.

Horn, J. L. & Cattell, R. B. "Refinement and test of the theory of fluid and crystallized general ingelligences". Journal of Educational Psychology, 57: 253-70, 1966.

Horn, J. L. & Cattell, R. B. "Age Differences in Fluid and Crystallized Intelligences." Acta Psychologica, 26: 107-29, 1967.

Howes, D. On the relation between the probability of a word as an association and in general linguistic usage. In L. A. Jakobovits and M. S. Miron (Eds.), Readings in the psychology of language. New Jersey: Prentice-Hall, 1967.

Hulicka, Irene. Age differences in retention as a function of interference. Journal of Gerontology, 22: 180-84, 1967.

Hulicka, Irene, Sterns, H. & Grossman, J. Age-group comparisons of paired-associate learning as a function of paced and self-paced association and responsive times. Journal of Gerontology, 72: 274-80, 1967.

Hunt, J. McV. Intelligence and Experience. New York: The Ronald Press, 1961.

Inglis, J. Immediate memory, age and brain functions. In A. T. Welford and J. E. Birren (Eds.), Chapter 7, 1965.

Jackson, D. N. Manual for the Personality Research Form. Goshen, New York: Research Psychologists Press, 1967.

Jackson, D. N. Multimethod factor analysis in the evaluation of convergent and discriminant validity. Psychological Bulletin, 72: 30-49, 1969.

Jackson, D. N. Jackson Personality Inventory. Copyright Douglas N. Jackson, University of Western Ontario, London, Ontario, 1970 (a).

Jackson, D. N. A sequential system for personality scale development. In Current Topics in Clinical and Community Psychology, Vol. 2. New York: Academic Press, 1970. (b)

Jackson, D. N., Hourany, L. J. & Vidmar, N. A. A four-dimensional interpretation of risk-taking. Journal of Personality, 40: 483-501, 1972.

Jacobson, L. J., Elenewski, J. J., Lordahl, D. S. and Liroff, J. H. Role of creativity and intelligence in conceptualization. Journal of Personality and Social Psychology, 10: 431-6, 1968.

Jamieson, G. H. Prior learning and response flexibility in two age groups. Journal of Gerontology, 24: 179-84, 1960.

Jarvik, L. F. & Cohen, D. A biobehavioral approach to intellectual changes with aging. In C. Eisdorfer and M. P. Lawton (Eds.), pp. 220-80, 1973.

Jensen, A. R. How much can we boost IQ and scholastic achievement? Harvard Educational Review, 39: 1-123, 1969.

Johnson, D. M. A Systematic Introduction to the Psychology of Thinking. New York: McGraw-Hill, 1972.

Jones, H. E. Intelligence and problem-solving. In J. E. Birren (Ed.) Handbook of Aging and the Individual. Chicago: University of Chicago Press, 1959.

Kalish, R. A.. Late Adulthood. Monterey: Brooks/Cole, 1975.

Kaplan, I. Foundations of human behavior. New York: Harper & Row, 1965.

Keevil-Rogers, P. & Schnore, M. M. Short-term memory as a function of age in persons of above average intelligence. Journal of Gerontology, 24: 184-8, 1969.

Kelley, T. I. Essential traits of mental life. Cambridge: Harvard University Press, 1935.

Kent, D. P., Kastenbaum, R. & Sherwood, S. (Eds.), Research, Planning, and Action For The Elderly. New York: Behavioral Publication, 1972.

Kimmel, D. C. Adulthood and Aging. New York: Wiley, 1974.

Kogan, N. Creativity and cognitive style: A life span perspective. In P. B. Baltes and K. W. Schaie (Eds.), Life Span Developmental Psychology. New York: Academic Press, Chap. 7, 1973.

Korchin, S. J. & Basowitz, H. Age differences in verbal learning. Journal of Abnormal and Social Psychology, 54: 64-9, 1957.

Krech, D., Crutchfield, R. S. & Livson, N. Elements of Psychology, 2nd ed., New York: Knopf, 1969.

Kucera, H. and Francis, W. N. Computational analysis of present-day American English. Providence, Rhode Island: Brown University Press, 1967.

Lawton, M. P. Assessing the competence of older people. In D. P. Kent, R. Kastenbaum, and S. Sherwood (Eds.), Chap. 10, 1972.

Lieberman, M. A. Institutionalization of the aged: Effects on behavior. Journal of Gerontology, 24: 330-40, 1969.

Lorge, I. & Thorndike, E. L. Semantic count of English Words. New York: Columbia University Press, Volumes 1-6, 1938.

Lowenthal, M. F. & Chiriboga, D. Social stress and adaptation: Toward a life-course perspective. In C. Eisdorfer and M. P. Lawton (Eds.), pp. 281-310, 1973.

McClelland, D. C. Testing for competence rather than for "intelligence." American Psychologist, 28: 1-14, 1973.

McNanamy, Eve W. Effects of novelty on time judgment by young and aged persons. Journal of Gerontology, 23: 41-4, 1968.

Mednick, S. A. The association basis of the creature process. Psychological Review, 69: 220-232, 1962.

Mednick, S. A. & Mednick, M. T. Examiner's Manual: Remote Associates Test. Boston: Houghton Mifflin, 1967.

Miles, W. R. The Stanford University studies of Later Maturity. In Welford A. T. & J. E. Birren, (Eds.), 1965.

Miner, J. B. On the use of a short vocabulary test to measure general intelligence. Journal of Educational Psychology, 52: 157-160, 1961.

Monge, R. H. Learning in the Adult Years--Set or Rigidity. Human Development, 12: 131-140, 1969.

Montagu, A. (Ed.), Race and IQ. New York: Oxford University Press, 1975.

Munnichs, J. M. Some problems of research about the intelligence of the aged. Gewein, 13: 268-279, 1965.

Nardi, A. H. Person-perception research and the perception of life-span development. In D. P. Kent, R. Kastenbaum, S. Sherwood (Eds.), Chap. 12, 1972.

Neugarten, B. L. Personality changes in later life: A developmental perspective. In C. Eisdorfer and M. P. Lawton (Eds), pp. 311-335, 1973.

Piaget, J. The Psychology of Intelligence, 1947 (Translated by M. Piercy & D. E. Berlyne, Otowa, New Jersey: Littlefield-Adams, 1966).

Pressey, S. L. & Pressey, A. D. Genius at 80 and other oldsters. The Gerontologist, 7: 193-7, 1967.

Rabbitt, P. "Age and the use of structure in transmitted information". In G. A. Talland, (Ed.) Human Aging and Behavior. New York: Academic Press, 1968.

Reichard, S., Livson, F., & Peterson, P. G. Aging and personality: A study of 87 older men. London: Wiley, 1962.

Riegel, K. F. A study of verbal achievement of older persons. Journal of Gerontology, 14: 453-6, 1959.

Riegel, K. F. Speed of performance as a function of age and set: A review of issues and data. In A. T. Welford and J. E. Birren (Eds.), Chap. 9, 1965.

Riegel, K. F. Riegel, Ruth M., & Levine, R. S. An analysis of associative behavior and creativity. Journal of Personality and Social Psychology, 4: 50-6, 1966.

Riegel, K. F., Riegel, Ruth., & Meyer, G. The prediction of retest resisters in research on aging. Journal of Gerontology, 23: 370-4, 1968.

Roget's International Thesaurus, 3rd Ed., New York: Thomas Y. Growell, 1962.

Rokeach, M. Long range experimental modification of values, attitudes and behavior. American Psychologist, 26: 453-9, 1971.

Rosenbaum, M. E., Arenson, S. J. & Panman, R. A. Training and instructions in the facilitation of originality. In C. P. Duncan (Ed.), Thinking: Current experimental studies. New York: J. P. Lippincott, pp. 214-223, 1967.

Rozeboom, W. W. Conditioned generalization, cognitive set, and the structure of human learning. Journal of Verbal Learning and Verbal Behavior, 6: 491-500, 1967.

Savage, R. D. & Bolton, N. A factor analysis of learning impairment and intellectual impairment in the elderly. The Journal of Genetic Psychology, 113: 177-82, 1968.

Savage, R. D. & Britton, P. G. The factorial structure of the WAIS in an aged sample. Journal of Gerontology, 23: 183-6, 1968.

Schaie, K. W. Developmental processes and aging. In C. Eisdorfer and M. P. Lawton (Eds), pp. 151-6, 1973.

Schaie, K. W., Labouvie, G. V., & Buech, B. U. Generational and cohort-specific differences in adult cognitive functioning: A fourteen-year study of independent samples. Developmental Psychology, 9: 151-166, 1973.

Schaie, K. W. & Strother, C. R. A cross-sequential study of age changes in cognitive behavior (1968). In P. S. Sears (Ed.), Intellectual Development. Toronto: John Wiley and Sons, 1971.

Szafran, J. Decision processes and aging. In A. T. Welford and J. E. Birren (Eds.), Behavior, aging, and the nervous system, Cahpter 2, 1965.

Taft, R. & Rossiter, J. R. The Remote Associates Test: Divergent of convergent thinking? Psychological Reports, 19: 313-14, 1966.

Taub, H. A. Aging and free recall. Journal of Gerontology, 23: 466-9, 1968.

Trembly, D. & O'Connor, J. Growth and decline of natural and acquired intellectual characteristics. Journal of Gerontology, 21: 9-12, 1966.

Tyler, Leona. The Psychology of Human Differences. New York: Merideth Publishing, 1965.

Vick, O. C. & Jackson, D. N. Cognitive styles in the schematizing process: A critical evaluation. Educational and Psychological Measurement, 27: 267-86, 1967.

Wallach, M. A. & Kogan, N. Modes of thinking in young children. New York: Holt, Rinehart & Winston, 1965.

Warren, T. P., & Davis, G. A. Distant associates and creativity: A new children's test. Proceedings, 78th Annual Convention, APA. p. 615, 1970.

Webster's New Twentieth Century Dictionary, 2nd Ed., Unabridged, New York: World Publishing, 1968.

Wechsler, D. The Measurement and Appraisal of Adult Intelligence. Baltimore: Williams and Wilkins, 1958.

Welford, A. T. & J. E. Birren (Eds.). Behavior, aging, and the nervous system. Springfield, Illinois: Charles C. Thomas, 1965.

Wetherick, N. E. Structure and content in concept attainment: A Full Scale Study. British Journal of Psychology, 60: 347-350, 1969.

Wilson, R. C., Christensen, P. R., Merrifield, P. R. & Guilford, J. P. Manual for the Alternate Uses Test. Beverly Hills: Sheridan Supply Co., 1960.

Zimmerman, I. L. & Woo-Sam, J. M. Clinical Interpretation of the Wechsler Adult Intelligence Scale. New York: Grune & Stratton, 1973.

APPENDIX 1

AGE GROUPINGS

Research on aging is in the fortunate position of having fairly close agreement between the research parameters defining age groups and conventional useage. It is fortunate because the definitions a lay person uses affects his attitude toward himself and behavior toward others (Anderson, 1967; Carp, 1969a). Cameron (1969) reports findings regarding these (usually implicit) age groupings. He asked his respondents (N=571), ages 11-80 what they though of when they heard or used the words (e.g.) "young adult"? The categories were: young adult--ages 18-25; middle age--40-55; old--65-80; and aged--80 and older.

There is general conformity to these parameters in the professional literature, although some experimenters might include 50 year olds with their "old" group (e.g., Gilberstadt, 1968; Riegel, 1965) and some may select some arbitrary age such as 40 at which to separate their groups (e.g., Jamieson, 1969). Selection of sample age limits can serve to either sharpen or attenuate results. For this paper age reference will be in respect to the following categories unless otherwise noted: young adult--ages 18-25; middle age--ages 30-59; old age--ages 60 and older.

APPENDIX 2

INDEX OF ASSOCIATIVE FLUENCY: RATIONALE AND PILOT STUDY

The rationale behind the development of the IAF was to provide a format wherein word power and the need to shift set were combined. It was hoped that the measurement of these relatively independent skills could be successfully integrated into one test. The instructions were designed to facilitate response alternation across grammatical classes. The stem words were selected from a list collected by surveying the dictionary for words with the largest number of varied meanings. For the response alternatives, both correct and incorrect, a number of criteria governed selection. First, the range of difficulty was to go from easy (very common) to hard (infrequent). The easy-to-hard dimension was equated with word frequency which was determined by reference to both the Lorge-Thorndike (1938) and the Kucera-Francis (1967) counts. Second, words that sound alike, but which were not necessarily synonyms, were included (e.g., plant-pliant). Third, words were selected that were structurally alike (e.g., rive, rave, raze). Fourth, at least one antonym was included among the incorrect alternatives. Fifth, words which were only nouns or only verbs were included among the alternatives. These criteria combined with the instructions to serve two

functions. The instructions prompted the person to expect more than one correct response and to be open to many kinds of meanings. Then words selected by the above five criteria present the person with a number of intriguing distractors which did not fit the instructions.

A pilot study was conducted with a sample (N=28) drawn from the Psychology Department, University of Western Ontario. The very high correlation with education of .81 was considered attributable to the selective nature of the sample, which was made up of secretaries, undergraduates, and candidates for graduate degrees. The sample statistics are:

Source	N	Range	Mean
Ph.D.'s	11	70-115	87
M.A.'s	7	60-95	74
Undergrad's	4	45-89	62
Secretaries	6	40-72	51

The scale was item-analyzed and correlations were computed between the score for each "item" (stem word) and the total score. These are slightly inflated because the score for each item was not deducted from the total score prior to the computations. Positive manifold was obtained with the average correlation being .64. These data along with the item means and variances (mean total = 72.38, variance = 404.43) are presented in Table A.

TABLE A

ITEM STATISTICS FOR THE IAF PILOT STUDY

Item	Possible #Correct	Mean	Variance	Correlation with total
abet	5	3.14	2.21	.47
abide	4	2.14	1.07	.18
bank	3	1.67	.89	.66
blow	7	3.05	1.47	.80
cast	7	2.10	1.04	.61
check	5	3.05	.81	.48
dash	4	1.90	.56	.68
draw	6	3.29	2.01	.89
fade	3	1.29	.30	.38
fine	4	2.95	1.09	.73
give	7	3.38	1.47	.84
hold	6	3.00	1.71	.69
issue	5	2.48	1.01	.71
key	4	2.57	1.01	.67
light	3	1.86	.50	.58
mark	5	2.52	1.49	.79
near	7	3.95	2.05	.81
pack	5	2.81	.82	.64
plug	4	1.62	.90	.67
range	3	1.10	.94	.60
report	7	2.95	1.28	.64
set	7	2.76	1.80	.77
shoot	5	2.95	1.19	.48
take	7	2.33	.98	.76
tear	7	2.57	1.20	.64
unite	5	2.71	.78	.43
upset	6	2.76	1.23	.77
work	5	2.43	.91	.68
zany	3	1.00	1.33	.51

APPENDIX 3

INDEX OF ASSOCIATIVE FLUENCY

NAME:_____ OCCUPATION:_____

AGE:_____ EDUCATION:_____ SEX:_____ DATE:_____

This test consists of a word followed by ten words which may or may not be similar in meaning. Any of the ten words can have the same meaning as the first word, and in every case, more than one answer is acceptable. You are to circle all of the associates both common and unusual. For example:

hard: size firm fluent obdurate franchise
bruise luck plenish arduous difficult

You would circle firm, obdurate, arduous, and difficult as they all mean hard in one sense or another. The choices must mean the same as the first word, but many meanings are included. When you are in doubt when making a comparison, reverse the comparison. That is, ask yourself the meaning of obdurate, for example. It does mean hard. However, even though hard and luck "go together", luck does not have as one of its meanings "hard".

While it is to your advantage to not omit any possibilities, and while first impressions are frequently correct, you should avoid wild guesses as they will be counted against the score.

Take as much time as you need to finish.

FORM A

abet:	inside	incite	foster	further	natural
	nuture	get	give	hope	help
abide:	flower	leave	wait	inhabit	endure
	evergreen	soul	bear	hide	look
bank:	hasp	heap	hole	spread	wealth
	lower	bestow	sell	store	cushion
cast:	mold	dash	nature	seize	arrive
	deposit	matrix	drop	lift	squint
blow:	stick	hit	toot	inspire	baffle
	flee	dissipate	misfortune	lam	pant
check:	king	count	test	relieve	curb
	post	police	arrest	confirm	languor
dash:	trace	cold	lethargic	rush	disappoint
	intrusion	hunger	sedate	digest	dull
draw:	tie	match	flame	neglect	describe
	lure	erase	inhale	condone	deduce
fade:	keen	cole	shrill	fervent	recede
	message	flag	distort	cover	eager

fine:	reward	course	keen	vulgar	minute
	dainty	metric	precise	porous	dubious
give:	plant	pliant	present	omit	move
	endow	accept	iterate	allow	proffer
hold:	grasp	restore	sale	fractious	purchase
	power	schism	contain	occupy	refrain
issue:	inherit	cause	apostate	effect	parent
	topic	monthly	shift	publication	flow
key:	island	pitch	resin	hole	clue
	index	fasten	latch	rise	skate
light:	heavy	obscure	intense	blight	heat
	idyll	ignite	young	airy	land
mark:	rate	neglect	define	slight	accent
	ideal	note	plague	dupe	prime
near:	close	about	impending	adjacent	detached
	intimate	continguous	literal	catch	reticent
pack:	pad	sparse	opine	flour	load
	pan	isolate	flock	gather	carry

plug:	form	depart	expose	irritate	nag
	goad	goat	chew	block	boost
range:	mix	vamp	rigid	rank	enact
	print	align	drama	home	solve
report:	acquit	opaque	sound	answer	critique
	award	recite	transfer	notice	rumor
set:	fix	class	etch	obscure	inform
	end	assign	trend	impose	clique
shoot:	diverge	recoil	dart	grow	leap
	jail	execute	project	ricochet	emit
take:	bestow	entail	exude	convey	infect
	eat	catch	deduce	deem	steal
tear:	mend	drop	rip	cleave	rage
	rive	rave	raze	ramp	hewn
unite:	sunder	identify	portion	join	discorp
	converge	votive	meet	own	wed
upset:	discompose	decay	revolution	quell	capsize
	denial	oppose	nerve	distress	sense

work:	create	hard	operate	radiate	manage
	arrange	strike	radar	resolve	accomplish
zany:	grimace	graceless	ingenuous	sham	vulgar
	ninny	jester	wag	rash	trite

APPENDIX 4

DETERMINATION OF WORD FREQUENCY

When word frequency is an important variable in the research, reference is usually made to the Thorndike-Lorge book, The Teacher's Word Book of 30,000 Words (1944), which is based on a count of twenty million words. However, it has a number of shortcomings which could interfere with the intent of the present research. First, only frequency is considered, no account is taken of meaning, that is, whether a word is used as a verb or noun, and so on. Second, the frequency is not given for all words--only category designations are given for the very frequent words. Third, the nature of the sample, which included the Bible and English classics, led to a high proportion of literary and learned words (e.g., pard, pistillate, ibid). Fourth, the entries do not include recent words while including words seldom seen today. Fifth, it includes proper names, the inclusion of which may change the rank-ordering of words.

For the above reasons other sources were appraised. Included among these were:

Horn, E. A basic writing vocabulary. (University of Iowa Monographs in Education, #4) Iowa City, 1926.

Kucera, H. & Francis, W. N. Computational analysis of present-day American English. Providence, Rhode Island: Brown University Press, 1967.

Lorge, I. The semantic count of the 570 commonest English words. New York: Columbia Teacher's College, 1949.

Lorge, I. & Thorndike, E. L. A semantic count of English words. New York: Columbia University, 1938. (6 volumes available on loan from Ohio State University.)

Roberts, A. H. A statistical linguistic analysis of American English. The Hague: Mouton, 1965.

West, M. A general service list of English words with semantic frequencies, 2nd Ed. London: Longmans, 1953.

West, M. International Readers Dictionary. London: Longmans, 1965.

Since semantic counts were desired, the Lorge-Thorndike (L-T, 1938) counts were used. However, a number of entries are estimated in this count as they come from different sample sizes, and these estimates give widely varying counts. For example, the original count (A) in 1933 drew on 13 different sources, a second count (B) which drew on 16 sources and which, for some words, was combined into a total count (C) of 29 sources (approximately five million words). West (1953) uses 2,000 words from the L-T count and provides a more detailed breakdown of the number of times the words occurred in each semantic category (e.g., as a noun, verb, and so on). The Lorge (1949) listing of the 570 commonest words also draws on the earlier L-T count. Since there were discrepancies between the L-T, the West, and the Lorge counts (or estimates based on some actual counts) an effort was made to estimate the frequency of all of the words used in the Index of Associative Fluency as though each frequency count was derived from a sample of five million words. If the actual count was based on 13 sources it was multiplied by 2.22 (29/16, the uncounted portion of the total count of 29 sources). If 16 sources had been used, the count was

multiplied by 1.82 (29/13). Example comparisons for two words are:

Word	Source	Estimate or Actual	Count
cover	L-T	Actual (16)	573
	Lorge	Estimate (29)	1240
	West	Estimate	1144
	Hourany	Estimate	1043
sound	I-T	Actual (13)	536
	Lorge	Estimate (29)	1134
	West	Estimate	1072
	Hourany	Estimate	1190

It was acknowledged (Lorge, 1949) that many clerical errors occurred in the tabulating of the words. For example, words which were marked as nouns in the 1938 count were marked as verbs in the 1949 count.

A more recent frequency count (Kucera & Francis, 1967) was obtained. The K-F count was derived from a sample of 1,014,232 "natural language" words drawn from 15 categories of written or printed material ranging from scientific writings, to general fiction, to personal letters.

The IAF words were scored for the percent they occurred as verbs, nouns, or adjectives, for the frequency with which they occurred as a verb, noun, or adjective, and for the total frequency according to the L-T count. The words were also scored according to the K-F count and scored for the selection frequency for this study. The results were correlated and are presented in Table B. The K-F frequencies and the selection frequencies for this sample.

(N=328) correlated .97. Therefore, the K-F counts are considered fairly representative of current selection frequencies (i.e., word familiarity) and the L-T and the sample frequencies are not further analyzed.

TABLE B

CORRELATIONS FOR THE SEMANTIC FREQUENCIES*

Category**	2	3	4	5	6	7	8	9
1	-86	-44	25	-36	-19	-04	-09	-04
2		-35	-37	32	-16	-07	-07	-10
3			-08	-21	23	-09	00	-02
4				18	55	85	67	71
5					08	77	55	52
6						78	80	72
7							88	85
8								97

*Decimals are omitted for clarity
**1=L-T percentage as a Verb
2=L-T percentage as a Noun
3=L-T percentage as an Adjective
4=K-F frequency as a Verb
5=K-F frequency as a Noun
6=K-F frequency as an Adjective
7=L-T total frequency
8=K-F total frequency
9=Frequency for this sample

APPENDIX 5

TABLE C

NUMBER OF SUBJECTS IN EACH TEST SEQUENCE

FOR EXPERIMENTS I AND II

EI Groups	Sequence*	Young N's	Middle N's	Old N
A	1234	5	6	7
B	1324	21	34	27
C	2314	4	6	5
D	2342	9	-	6
E	2143	4	-	7
F	3124	20	27	21
G	4321	-	16	-
EII Groups				
A	1324	33	37	39
B	3124	47	49	41

Transfer Analyses**	N's for EI		N's for EII	
	1	2	1	2
Young	21	20	33	38
Middle	34	27	37	49
Old	27	21	39	41

*1=QWT, 2=AU, 3=IAF, 4=RAT

**1=When the IAF followed the dependent variable.

2=When the IAF preceeded the dependent variable.

APPENDIX 6

PERSONALITY QUESTIONNAIRE

The questionnaire consists of eight scales with 16 questions each. The subject is to decide whether the statement is true or is descriptive of him, or whether it is false or not descriptive. The following is an outline of trait descriptions, illustrative of people who get high scores on the individual trait scales. A low scorer would have opposite characteristics.

1. Conformance

Respects authority and enjoys clearly defined rules and regulations, cooperates fully with leadership, accepts and benefits from criticism.

2. Value Orthodoxy

Traditional, moralistic, conventional, values traditional customs and beliefs, values may be seen by others as old fashioned, takes a rather conservative view regarding contemporary standards of behavior.

3. Change

Variable, flexible, restless, likes new and different experiences, dislikes routine and avoids it, may change opinions or values in different circumstances.

4. Innovation

Prefers the new to the old, seeks original solutions to old problems, believes everything can be improved, would rather

try something experimental and different than tried and true, would rather invent than develop.

5. Autonomy

Trys to break away from restraints, confinement or restrictions of any kind, enjoys being unattached, free, not tied to people, places, or obligations, may be rebellious when faced with restraints.

6. Assertiveness

Attempts to control the environment, attempts to influence or direct other people, forceful, decisive, authoritative.

7. Impulsiveness

Tends to act or speak on the "spur of the moment" and without deliberation of hesitation, spontaneous, hasty, impetuous, uninhibited.

8. Achievement

Aspires to accomplish difficult tasks, maintains high standards and is willing to work toward distant goals, responds positively to competition, willing to put forth effort to attain excellence.

APPENDIX 7

TABLE D

REPETITIONS IN EI AND EII

The following entries represent the scores for the subjects who participated in both EI and EII. The first two digits refer to the group number and the second two refer to age

Code No.	1970 Scores					1972 Scores				
	IAF	ER.	QWT	RAT	AU	IAF	ER.	QWT	RAT	AU
253411207	77	6	76	18	23	82	1	54	12	18
253511277	37	3	45	13	19	84	7	55	19	31
254810810	61	3	78	21	14	76	2	80	27	24
255611202	80	0	67	18	27	103	2	68	15	27
312511010	62	3	81	13	19	65	1	65	19	18
312710909	82	11	76	16	14	49	3	63	17	20

Members of Group 25 received the IAF first, while members of Group 31 received the QWT first.

APPENDIX 8

TABLE E

CORRELATIONS BETWEEN THE IAF RAW AND WEIGHTED MEASURES AND EACH OF THE COMPARISON MEASURES*

Age Groups	Experiment I					
	QWT		RAT		AU	
	A**	B	A	B	A	B
17-26	55	46	54	56	34	36
30-59	50	64	29	30	33	33
60-89	55	57	24	23	15	16
17-89	47	50	32	32	21	21
	Experiment II					
	QWT		RAT		AU	
17-26	55	58	54	53	34	38
30-59	54	56	29	31	24	27
60-89	34	37	33	33	15	17
17-89	47	50	32	32	21	21
	UT Fluency		UT Flexibility		Education	
17-26	33	33	14	11	21	19
30-59	23	24	04	05	15	17
60-89	20	21	29	27	29	27
17-89	21	22	10	11	16	17

*Decimals have been omitted.
**A=IAF Raw score; B=IAF Weighted score.

APPENDIX 9

GRAMMATICAL ANALYSIS

A comparison was made of the relative proportion of times Ss at each age level selected items in each of the word categories. The results are reported in Table G. The categories were combinations of grammatical classes. For example, words designated by Webster's New Twentieth Century Dictionary (1968) as nouns, and words designated as being both a noun and an adjective were collapsed into Category 1, and so on. The words were combined into these categories because of the infrequency of some classes. For instance, only two words were designated as adverbs so they were combined with the verbs to form Category 2. The cell entries are the percentage of times each age group selected a word from a given category. The age groups were: young, 17-26 (N=128); middle, 30 59 (N=95); old, 60-91 (N= 105). In no instance was there greater than a four percent difference between young, middle, and old age groups in their word selection patterns. For IAF words keyed either correct or incorrect, there were no differential tendencies between age groups to select or omit words from any grammatical category.

TABLE F

SELECTION FREQUENCIES IN RELATION TO GRAMMATICAL CATEGORY

Category*	Grammatical Class**	Age Group	%correctly selecting key Correct	% not selecting key Correct	%correctly selecting key False	%incorrectly selecting key False
1	Noun N-Adj.	Young	14	55	29	02
		Middle	16	54	27	03
		Old	16	51	28	05
2	Verb Adverb	Young	27	39	32	02
		Middle	27	39	32	02
		Old	24	37	34	04
3	Adjective	Young	10	72	12	06
		Middle	11	72	12	05
		Old	08	71	14	06
4	V-Adv-Adj V-Adj.	Young	08	73	12	06
		Middle	11	74	09	06
		Old	11	69	09	11
5	N-V	Young	24	39	36	01
		Middle	24	38	35	02
		Old	25	36	35	04
6	N-V-Adj N-V-Adv N-V Adv-Adj	Young	24	55	18	03
		Middle	24	55	18	03
		Old	24	51	18	07

*The total number of times the 290 IAF words occurred in each category is: 1=30; 2=58; 3=22; 4=10; 5=139; 6=31.

**Grammatical class was determined by reference to Webster's New Twentieth Century Dictionary, 1968.

APPENDIX 10

EXPERIENCE AND SKILL

If the claim that experience "produces" a skill or somehow augments an innate propensity is viable, differences should appear in the relationship of the various measures and education. To examine this possibility the relevent correlations are presented in Table G. Education is commonly considered to provide a sharpening of the verbal skills. Indeed, the stereotype is so prevalent it is the source of numerous jokes. An examination of the correlations reveals, however, that a somewhat stronger relationship exists between education and a test (the AU) requiring varied uses for common objects, a skill which could reasonably be expected to come from experience. That an education-produced test-taking facility is not being measured is apparent from the higher correlations for the old in comparison to the young. The correlations between performance on these tests and actual school achievement (Grade as reported by each subject) yields no defineable pattern.

APPENDIX 10

TABLE G

CORRELATIONS FOR THE PERFORMANCE MEASURES AND EDUCATION*

Experiment I Age Group	Education IAF	QWT	RAT	AU
17-26	03	04	26	13
30-59	05	16	00	22
60-89	02	34	20	30
17-89 (All Ss)	13	20	35	43

Experiment II Age Group	Education IAF	QWT	RAT	AU
17-26	21	19	08	08
30-59	15	22	09	31
60-89	29	48	08	23
17-89 (All Ss)	14	19	03	35
	GRADE IAF	QWT	RAT	AU
17-26	23	23	17	06
3-59	14	19	11	12
60-89	19	36	21	29
17-89 (All Ss)	16	31	18	24

*Decimals are omitted for clarity.

APPENDIX H1

CONTROLLING FOR EDUCATION

There is some evidence that differences in educational attainment affect the age-ability relationship (Granick & Friedman, 1967). Since the level of formal education differs significantly for the sample of Ss used in this study, its effects were assessed by a covariance analysis. A comparison was made between an unmodified analysis of variance and an analysis controlling for education. The two analyses produced almost identical results for the QWT and the IAF. For the RAT and the AU there is some reduction in the magnitude of the age-related decrement in performance on these measures. These results are presented in Table H. It should be noted that the relationship between ability level and performance is not affected by differences in education. Even though there was a slight reduction in the age effect for the RAT and the AU, the basic findings for all of the measures is essentially unchanged.

APPENDIX 11

TABLE N

ANALYSIS OF VARIANCE CONTROLLING FOR EDUCATION*

	Analysis of Variance			Analysis of Covariance		
Variable	F	P	ω^2	F	P	ω^2
RAT - Age effect	70.2	.001	.31	47.8	.001	.24
Ability effect	10.3	.001	.06	9.1	.001	.06
Interaction	1.2					
AU - Age effect	60.1	.001	.27	31.2	.001	.17
Ability effect	11.4	.001	.07	9.4	.001	.07
Interaction	1.9			1.5		

*Equal N's of 23 per cell were used for this analysis.